murine coro cucurbita m haemophilus

bacillus su salmonella

homo sapien

homo sapien

rhizobium m pedobacter mus musculu

norv

rattus

```
SEQUENCE FROM N.A.
MEDLINE-90152677; PubMed=1968043;
Schuchman E.H., Jackson C.E., Desnick R.J.;
"Human arylsulfatase B: MOPAC cloning, nucleotide sequence of a full-length cDNA, and regions of amino acid identity with arylsulfatases A and C.";
                                                                                                                                                                                                                                                                                                                                                                                                                                             (N-acetylgalactosamine-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     K.; that is defective in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 1-104 FROM N.A.

SEQUENCE OF 1-104 FROM N.A.

MEDLINE=92028992; PubMed=1930244;

Litjens T., Morris C.P., Gibson G.J., Beckmann K.R., Hopwood J.J.;

Litjens T., Morris C.P., Gibson G.J., Beckmann R.R., Hopwood J.J.;

"Human N-acetylgalactosamine-4-sulphatase: protein maturation and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
MEDLINE-97184692; Pubmed-9032078;
Bond C.S., Clements P.R., Ashby S.J., Collyer C.A., Harrop S.J.,
Hopwood J.J., Guss J.M.;
"Structure of a human lysosomal sulfatase.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
MEDILINE-90153994; PubMed=2303452;
Peters C., Schmidt B., Rommerskirch W., Rupp K., Zuehlsdorf M., Vingron M., Meyer H.E., Pohlmann R., von Figura K.;
"Phylogenetic conservation of arylsulfatases. cDNA cloning and expression of human arylsulfatase B.";
Biol. Chem. 265:3374-3381(1990).
                                                 059288
P97333
P11225
09x442
0060674
0007833
                   P97874
Q52938
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Modaressi S., Rupp K., von Figura K., Peters C., "Structure of the human arylsulfatase B gene."; Biol. Chem. Hoppe-Seyler 374:327-335(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Schmidt B., Selmer T., Ingendoh A., von Figura "A novel amino acid modification in sulfatases
                                                                                                                                                                                                                                                                                                                                                                                 01-APR-1990 (Rel. 14, Created)
01-APR-1990 (Rel. 14, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
Arylsulfatase B precursor (EC 3.1.6.12) (ASB)
4-sulfatase) (GAS).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2-AMINO-3-OXOPROPIONIC ACID MODIFICATION SITE
                                                                                                                                                                                                                                                                                                                                                   533 AA
                                                                                                                                                                                                                                                         ALIGNMENTS
                                                                      NRP1_MOUSE
VGL2_CVMJH
ASO_CUCMA
HGPC_HAEIN
NRP1_RAT
GAK_RAT
RKP1_RHIME
CHAC_PEDHE
                                                                                                                                                              WAPA_BACSU
YEJM_SALTY
PTPJ_HUMAN
                                                                                                                                                JAK2_HUMAN
                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
MEDLINE-93332648; PubMed-7687847;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=95354208; PubMed=7628016;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ":nultiple sulfatase deficiency.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           isolation of genomic clones.";
Biochem. Int. 24:209-215(1991)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Genomics 6:149-158(1990).
                                                                                                                                                                                                                                                                                                                                                   STANDARD;
922
1305
538
700
923
1235
579
1132
2334
586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Cell 82:271-278(1995).
\begin{matrix} \mathbf{u} & \mathbf{u} 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID~9606;
                                                                                                                                                                                                                                                                                               ARSB_HUMAN
ID ARSB_HUMAN
AC P15848;
VT 01-*-
103.5
102.5
102.5
102
102
100.5
100.9
99
99
98.5
97.5
 (without alignments)
370.748 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            felis silve rattus norv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    mus musculu
klebsiella
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          mus musculu
rhizobium m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  oryza sativ
escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         strongyloce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            rattus norv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          hemicentrot
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            capra hircu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                volvox cart
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     chlamydomon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              haemophilus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               bacillus su
emericella
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Description
                                                                                                                                                                                                                                    MAPRGCAGHPPPPSPQACVC......VPVRYPPKDPRSNPRLNGGV
                                                                                                                        July 23, 2002, 19:01:44; Search time 57.44 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P34059
P50473
P08842
P51690
P15589
P25549
P50427
P50427
P15089
P50428
P50427
P1000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          P15586
Q10723
P31447
P14217
P51688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       069787
P44135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          P54496
Q9p4p9
P27934
P39401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               P54793
P51691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P77318
P50426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             P50430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          068800
                                    Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                              Potal number of hits satisfying chosen parameters:
               GenCore version 4.5
Copyright (c) 1993 - 2000 Comp
                                                                                                                                                                                                                                                                                                                              105224 seqs, 38719550 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ARSB_MOUSE
ARSF_HUMAN
ARS_PSEAE
GA6S_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STS_HUMAN
ARSE_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ARSB_HUMAN
ARSB_FELCA
ARSB_RAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ARSD_HUMAN
ASLA_ECOLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ARS_VOLCA
YIDJ_ECOLI
ARS_CHLRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       YDEN_ECOLI
GL6S_CAPHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AM3E_ORYSA
MDOB_ECOLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            YQGS_BACSU
HIS5_EMENI
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ARSA_HUMAN
ARSA_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GL6S_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SPHM_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BETC_RHIME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              rc46_haein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HUMAN
                                                                                       protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STS_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ARS_KLEAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ARS_HEMPU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ARS_STRPU
                                                                                                                                                                                                                                                                       BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IDS_
                                                                                                                                                                                                                                                                                                                                                                                                   Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                          US-09-495-823-7
3012
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SwissProt_40:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1426.5
1414.5
1333.3
1333.3
864
430.5
429
422.5
416.5
407.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score
                                                                                                                                                                                                Title:
Perfect score:
                                                                                                                                                                                                                                                                          Scoring table:
                                                                                       protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Database :
                                                                                                                                                                                                                                    Sequence:
                                                                                                                                                                                                                                                                                                                              Searched:
                                                                                                                           Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
```

```
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                     VARIANTS MPS-VI ARG-117; PRO-236 AND TYR-405.
MEDLINE-92197625; PubMed-1550123;
Jin W.-D., Jackson C.E., Desnick R.J., Schuchman E.H.;
"Mucopolysaccharidosis type VI: identification of three mutations in the arylsulfatase B gene of patients with the severe and mild phenotypes provides molecular evidence for genetic heterogeneity.";
Am. J. Hum. Genet. 50:795-800(1992).
                                                                                                                                        intermediate
                                                                                                                                                                                                                      Ł
                                                                                                                                       syndrome). An intermediat of valine for glycine at
                                                                                                                                                                                     VARIANTS MPS-VI MET-92; GLN-95; CYS-210; PRO-393 AND PRO-498. MEDLINE-96213747; PubMed-8651289; Littlens T. Brooks D.A., Peters C., Gibson G.J., Hopwood J.J.; Infentification, expression, and biochemical characterization of acceptigalactosamine-4-sulfatase mutations and relationship with clinical phenotype in MPS-VI patients."; Am. J. Hum. Genet. 58:1127-1134(1996).
                                                                                                                                                                                                                                                                   VARÍANTS MPS-VI TRP-152 AND GLN-160.
MEDLINE-94171224; Pubmed-8125475;
Voskoboeva E., Isbrandt D., von Figura K., Krasnopolskaya X.,
                                                                                                                                                                                                                                                                                                        "Four novel mutant alleles of the arylsulfatase B gene in tw
patients with intermediate form of mucopolysaccharidosis VI
(Maroteaux-Lamy syndrome).";
Hum, Genet. 93:259-264(1994).
                                                                                                        MEDLINE-92042029; PubMed-1718978;
Wicker G., Prill V., Brooks D., Glbson G., Hopwood J.,
von Figura K., Peters C.;
"Mucopolysaccharidosis VI (Maroteaux-Lamy syndrome). An
clinical phenotype caused by substitution of valine for
position 137 of arylsulfatase B.";
J. Biol. Chem. 266:21386-21391(1991).
                                                                                                VARIANT MPS-VI VAL-137, AND VARIANT MET-376.
Structure 5:277-289(1997)
                                                                                                                                                                                                                                                                                                 Peters
                                                                                                                                                                                                                                                           10]
```

```
PDB; its.,

( MIM; 253200; -.

( MIM; 272200; -.

R InterPro; IPR000917; Sulfatase.

)R PROSTE; PS000523; SulFATASE_1; 1.

DR PROSTE; PS000149; SulFATASE_2; 1.

KW Hydrolase; Signal; Glycoprotein; Lysosome; Mucopolysaccharldosis;

KW Disease mutation; Polymorphism; 30-printlad.).

FT 353 ARYLSULFATASE B.

1 7533 ARYLSULFATASE B.

1 77 147 POTENTIAL.

2-AMINO-3-OXOPROPIONIC ACID.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QAGEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQP 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PPGS----GAGASRPPHLVFLLADDLGWNDVGFHGSRIRTPHLDALAAGGVLLDNYYTQP 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLA 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               G -> V (IN MPS-VI; INTERMEDIATE FORM).
/FTIG4-MR_007297.
R -> W (IN MPS-VI; INTERMEDIATE FORM).
/FTIG4-VAR_007298.
R +> O (IN MPS-VI; INTERMEDIATE FORM).
/FTIG4-VAR_007299.
Y -> C (IN MPS-VI; MILD/INTERMEDIATE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      H -> P (IN MPS-VI; MILD/SEVERE FORM).
/FIId-VAR_007304.
C -> Y (IN MPS-VI; MILD FORM).
/FIId-VAR_007305.
L -> P (IN MPS-VI; MILD/SEVERE FORM).
/FIId-VAR_007306.
                                                                                                                                                                                                                                                                                                                                                                                                                                           -> Q (IN MPS-VI; MILD/SEVERE FORM).
                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                 . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            57;
                                                                                                                                                                                                                                                                                                                                                            (PROBABLE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              47.4%; Score 1426.5; DB 1; Length 533;
49.3%; Pred. No. 4.7e-105;
ive 81; Mismatches 149; Indels 57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEVERE FORM).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -> R (IN MPS-VI; SEVERE FORM).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /FTId-VAR_007300.
L -> P (IN MPS-VI; MILD FORM).
                                                                                                                                                                                                                                                                                                                                                                                                                 -> M (IN MPS-VI; MILD FORM).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -> M (IN REF. 3).
5983FB6911C4789A CRC64;
                                                                                                                                                                                                                                                                                                                                                              (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC. . . . N-LINKED (GLCNAC. . . .
                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      R (IN MPS-VI;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /FTId=VAR_007296.
G -> V (IN MPS-VI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /FTId=VAR_007302.
V -> M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /FTId=VAR_007301.
                                                                                                                                                                                                                                                                                                                                                                                                                              /FTId=VAR_007294
                                                                                                                                                                                                                                                                                                                                                                                                                                                         007295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /FTId=VAR 007303
                                                                                                                                                                                                                                                                                                                                                                                                                                                        FTIG-VAR
C -> R (IN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -^ W
JOINED.
JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .;
Έ
X72738; CAA51272.1; UX72739; CAA51272.1; UX72741; CAA51272.1; UX72741; CAA51272.1; UX72742; CAA51272.1; UX72777; AAB19988.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        59687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                             2521
1922
1922
1923
1988
2013
4014
902
903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                498
                                                                                                                                                                                                                                                                                                                                                                                                                                          95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     405
                                                                                 5078; A35078
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       533 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
Matches 279; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     405
                                                                                                                                                                                                                                                                                                                                                                                                                                          95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                376
                                                                                                                                                                                                                                                                                                                   DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                 VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                             VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARIANT
                                                      EMBL;
EMBL;
 EMBL;
EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

œ

CAA51272.1; -. CAA51272.1; JOINED.

EMBL; JO5225; AAA51784.1; -. EMBL; M32373; AAA51779.1; -. EMBL; X72735; CAA51272.1; -. EMBL; X72736; CAA51272.1; JOI

 $\sim$ 

```
39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
  P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    엄
                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                            415
                                                                                                                                                                                                               235
                                                                                                                                                                                                                                                                                                                416 VDILHNIDPIY-----TKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGNPGYS 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-FEB-1994 (Rel. 28, Last sequence update)
15-701-1998 (Rel. 36, Last annotation update)
Arylsulfatase B precursor (EC 3.1.6.12) (ASB) (N-acetylgalactosamine-
4-sulfatase) (G4S).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CDNA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       dermatan sulfate.
SUBUNIT: HOMODIMER.
SUBCELLULAR LOCATION: Lysosomal.
DISEASE: DEFECTS IN ARSB ARE THE CAUSE OF MAROTEAUX-LAMY
SYNDROME; ALSO KNOWN AS MUCOPOLYSACCHARIDOSIS TYPE VI (MPS-VI)
MPS-VI HAS BEEN DESCRIBED IN SIAMESE CATS.
SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
                                                              Genomics 14:403'411(1992).
-!- CATALYTIC ACTIVITY: Hydrolysis of the 4-sulfate groups of the acetyl-D-galactosamine 4-sulfate units of chondroitin sulfate
ICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLG
            MLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRA
                                              FYRKECMPTRRGFDTFFGSLLGSGDYYTHYKCD----SPGMCGYDLYENDNAAWDYDNG
                                                                                                 IYSTQMYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRSIININRRRYAA
                                                                                                                                                                                                 VGFVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPR
                                                                                                                                                                                                                                                                                                   DWVPPQSFSNLGPNRWHNERITSS--TGKSVWLFNITADPYERVDLSNRYPGIVKKLLRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TISSUE-Liver;

MEDLINE-93052342; PubMed-1427856;
Jackson C. E., Yuhki N., Desnick R.J., Haskins M.E., O'Brien S.J., Schuchman E.H.;
"Felline arylsulfatase B (ARSB): isolation and expression of the ci comparison with human ARSB, and gene localization to feline chromosome Al.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Felis silvestris catus (Cat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
                                                                                                                                                                                                                                                                                                                                                                                                                                         535 AA
                                                                                                                                                                                                                                                                                                                                                   LSQFNKTAVPVRYPPKDPRSNPRLNG 548
                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                   | ::| :|| :|| :|: | | | 502 LQFYHKHSVPVYFPAQDPRCDPRATG
                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-FEB-1994 (Rel. 28, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=9685;
                                                                                                                                                                                                                                                                                                                                                                                                                                         ARSB_FELCA
                                                                                                                                                                                                                                                                                                                                                                                                                          ARSB_FELCA
                                                                                                                                                                                                 356
                                                                                                                                                                                                                                                                                                   465
                                                                                                                                                                                                                                                                                                                                                  523
                        90
                                                                                                                        209
                                                                                                                                                 296
                                                                                                                                                                        569
121
                                                                                               236
                                                                                                                                                                                                                         329
                                                 181
                       8
                                                                     g
                                                                                               δ
                                                                                                                       g
                                                                                                                                                 ò
                                                                                                                                                                       q
                                                                                                                                                                                                 ò
                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ç
                                                 õ
```

```
289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        55 EGALLAQAGEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLE 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NYYVQPICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 175 GKWHLGFYRKECMPTRRGFDTFFGSLLGSGDYYTHYKC---DSPGM--CGYDLYENDNAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WDYDNGIYSTQMYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRSIININ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNGGQPTAGGSNWPLRGSKGTYW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       350 EGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      410 GLRSPRVDILHNIDPIY-----TKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    459 GNPGYSDWVPPQSFSNLGPNRWHNERITSS--TGKSVWLFNITADPYERVDLSNRYPGIV
                                                            HSSP; P15848; 1100.
InterPro: JPR00917; Sulfatase.
Pfam; PF0084; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
PROSITE; PS00149; SULFATASE_2; 1.
Hydrolase; Signal; Glycoprotein; Lysosome; Mucopolysaccharidosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 MAPRGCAGHPPPPSPQACVCPG-----KMLAMGALAGFWILCLLTYGYLSWGQALEEEE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MGRRGAASLPRGPSPRRPLLPGVLPLLLRLLLPSRPG---------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----AGAG------ADRPPHLVFVLADDLGWNDVSFHGSNIRTPHLDELAAGGVLLD
                                                                                                                                                                                                                                                                                                (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                          67;
                                                                                                                                                                                                                                                                                                                                                                                                                             Length 535;
                                                                                                                                                                                                                                                                                                                                                                                                                          47.0%; Score 1414.5; DB 1; Length
49.0%; Pred. No. 4.2e-104;
Live 82; Mismatches 143; Indels
                                                                                                                                                                                                                                                                                                                                                                             43A527886A9983C4 CRC64
                                                                                                                                                                                      POTENTIAL.
2-AMINO-3-OXOPROPIONIC
                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC. N-LINKED (GLCNAC. N-LINKED (GLCNAC. N-LINKED (GLCNAC. N-LINKED (GLCNAC. N-LINKED (GLCNAC.
                                                                                                                                                                                                                      (BY SIMILARITY)
                                                                                                                                                                                                                                BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 517 KKLLRRLSQFNKTAVPVRYPPKDPRSNPRLNG 548
send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ::|| || ::| :|| :|| :| :|| :| :|| EQLLSRLQFYHKHSVPVHFPAQDPRCDPRGTG
                                                                                                                                                                                                                                                                                                                                                                                 MΚ
                             EMBL; S48472; AAB23941.1; -.
                                                                                                                                                                                                                                  523
1157
1194
449
1190
2281
2293
428
428
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                       535
149
93
                                            PIR; A44475; A44475.
HSSP; P15848; 1FSU.
                                                                                                                                                                                                                                                                                                                                                                                 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                  119
123
183
407
190
281
293
428
460
                                                                                                                                                                                    149
93
                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Sim
Matches 280;
                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                      ACT_SITE
MOD_RES
                                                                                                                                                                                                                                    DISULFID
                                                                                                                                                                                                                                                                   DISULFID
                                                                                                                                                                                                                                                                                     DISULFID
                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
```

328

181

388

361

```
242 GGOTRSGGNNWPLRGRKGTLWEGGIRGAGFVASPLLKOKGVKSRELMHITDWLPTLVNLA 301
                                                                                                                                                                                                            EGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPIY------TKAKNGSWAAG 437
                                                                                                                                                                                                                                                                                                     362 HSAFNTSIHAGIRYKNWKLLTGYPGCGYWFPPPSQSNISEVPS-----VDSPTKTLWL 414
 PLDEKLLPQLLKDAGSSTHMVGKWHLGMYRKECLPTRRGFDTYFGYLLGSEDYYTHEACA 122
                                                                                                                                                                                                                              -YKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQILASHNPTKPIFLYIAYQAV
                                                                                         HSPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDN
                                                                                                                                                    GGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLISLA
                                                                                                                                                                                                                                                                          YGIWNTAIQSAIRVQHWKLLTGNPGYSDWVPPQSFSNLG--PNRWHNERITSSTGKSVWL
                                                                                                                                                                                                                                                                                                                                   FNITADPYERVDLSNRYPGIVKKLLRRLSQFNKTAVPVRYPPKDPRSNPRLNG 548
                                                                                                                                                                                                                                                                                                                                                    P50429;
                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 4
ARSB_MOUSE
                                                                                         569
                                                                                                                                                    329
                                                                                                                                                                                                              389
                                                                                                                                                                                                                                          302
                                                                                                                                                                                                                                                                          438
                                                                                                                                                                                                                                                                                                                                    496
                                                                                                                                                                                                                                                                                                                                                                 415
                                                          9
                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                 qq
                            ò
                                                                                         à
                                                                                                                   a
                                                                                                                                                    ò
                                                                                                                                                                              q
                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                               -I- CATALYTIC ACTIVITY: Hydrolysis of the 4-sulfate groups of the N-acetyl-D-galactosamine 4-sulfate units of chondroitin sulfate and dermatan sulfate.
-I- SUBUNIT: HOMODIMER (BY SIMILARITY).
-I- SUBCELLULAR LOCATION: Lysosomal.
-I- DISEASE: DEFECTS IN ARSB ARE THE CAUSE OF MAROTEAUX-LAMY SYNDROME; ALSO KNOWN AS MUCOPOLYSACCHARIDOSIS TYPE VI (MPS-VI).
-I- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
                                                                                                                                                    Craniata; Vertebrata; Euteleostomi;
Sciurognathi; Muridae; Murinae; Rattus.
                                                                                                                                                                                                                                                                        "Mucopolysaccharidosis type VI in rats: isolation of cDNAs encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2-AMINO-3-OXOPROPIONIC ACID (BY SIMILARITY).
                                                       01-0CT-1996 (Rel. 34, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
Arylsulfatase B (EC 3.1.6.12) (ASB) (N-acetylgalactosamine-
                                                                                                                                                                                                                                                                                      arylsulfatase B, chromosomal localization of the gene, and
identification of the mutation.";
Genomics 29:582-587(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Aydrolase; Glycoprotein; Lysosome; Mucopolysaccharidosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4E114C923A24AF8F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
               473 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR000917; Sulfatase.
Pfam; PF00084; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; FALSE_NEG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                              STRAIN-SPRAGUE-DAWLEY; TISSUE-Liver;
                                                                                                                                                                                                                                          MEDLINE-96121368; PubMed-8575749;
                                          01-OCT-1996 (Rel. 34, Created)
                                                                                                                                                  Eukaryota; Metazoa; Chordata;
Mammalla; Eutheria; Rodentia;
                                                                                        Arylsulfatase B (EC 3.1.6.12)
4-sulfatase) (G4S) (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               53320 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; D49434; BAA08412.1; -.
               STANDARD;
                                                                                                                                    Rattus norvegicus (Rat).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            461
95
132
387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   231
355
366
398
473 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PSE
                                                                                                                                                                                              [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                NCBI_TaxID=10116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             87
57
61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P15848
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                            Kunieda
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NON_TER
ARSB_RAT
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseéisb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 34-285 FROM N.A.

MEDLINE=96323200; PubMed=8710849;

MEDLINE=96323200; Schmidt P., Hafner A., McLoghlin D.B.,

Schmahl W., Hess B., von Figura K., Peters C.W.B.;

"Targeted disruption of the arylsulfatase B gene results in mice resembling the phenotype of mucopolysaccharidosis VI.";

Proc. Natl. Acad. Sci. U.S.A. 93:8214-8219(1996).

-! CATALYTIC ACTIVITY: Hydrolysis of the 4-sulfate groups of the N-acetyl-D-galactosamine 4-sulfate units of chondroitin sulfate and
                                                                                                                                                                                                                                                                                    Grompe M., Pieretti M., Caskey C.T., Ballabio A.;
"The sulfatase gene family: cross-species PCR cloning using the MOPAC
                                                                                                                                                                         Euteleostomi;
Murinae; Mus
                                01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Arylsulfatase B (EC 3.1.6.12) (ASB) (N-acetylgalactosamine-4-sulfatase) (G4S) (Fragments).
                                                                                                                                                                         Craniata; Vertebrata; E
Sciurognathi; Muridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -i- SUBCELLULAR LOCATION: Lysosomal.
-i- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
285 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HOMODIMER (BY SIMILARITY).
PRT;
                                                                                                                                                                                                                                                SEQUENCE OF 1-33 FROM N.A.
MEDLINE-92241876; Pubmed-1572648;
                                                                                                                                                                       Chordata; (Rodentia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; M82877; AAA37261.1; -. EMBL; X92096; CAA63067.1; -.
                                                                                                                                                                                                                                                                                                                                               Genomics 12:755-760(1992).
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             dermatan sulfate.
SUBUNIT: HOMODIME
                                                                                                                                                     Mus musculus (Mouse)
                                                                                                                                                                       Eukaryota; Metazoa;
Mammalia; Eutheria;
                                                                                                                                                                                                                NCBI_TaxID=10090;
                                                                                                                                 ARSB OR AS1
ARSB_MOUSE
                                                                                                                                                                                                                                                                                                                             technique.
```

5.

Gaps

26;

Indels

44.3%; Score 1333; DB 1; Length 473; 53.9%; Pred. No. 9.6e-98;

Mismatches

10;

Conservative

Similarity

Mc Local 5-255;

Best Loca Matches

ð 8

Query Match

PLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDTFFGSLLGSGDYYTH---- 209

154

õ

94 HGSEIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCL 153

S

```
MEDLINE=95236447; PubMed=7720070;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hydrolase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                400
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                            .;
m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        151 NCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDTFFGSLLGSGDYYTHY 210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                211 KCD-----SPGMCGYDLYENDNAAMDYDNGIYSTQMYTQRVQQILASHNPTKPIFLYIAY 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    91 VGYHGSEIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITGKYQIHTGLQHSIIRPTQP 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLI 385
                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .
G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      the
                                                                                                                                                                                                                                                                                                                                                                                                                                   QAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                    AC. . .) (POTENTIAL).
AC. . .) (POTENTIAL).
AC. . .) (POTENTIAL).
AC. . .) (POTENTIAL).
                                                                                                                                                                                                     (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUE-Fetal brain;
MEDLINE=97336043; PubMed-9192838;
MEDLINE=97336043; PubMed-9192838;
Puca A.A. Zollo M., Repetto M., Andolfi G., Guffanti A., Simon Ballabio A., Franco B.,
"Identification by shotgun sequencing, genomic organization, and functional analysis of a fourth arylsulfatase gene (ARSF) from t
                                                                                                                                                                                                                                                                                                                                                                            52;
                                                                                                                                                                                                                                                                                                                                   28.7%; Score 864; DB 1; Length 285; 49.8%; Pred. No. 5.3e-61; ive 45; Mismatches 68; Indels
                                                                                                                                                                                                                                                                                MW; 36BA5B98EA0770C9 CRC64;
                                                                                                                                                                                N-LINKED (GLCNAC. N-LINKED (GLCNAC. N-LINKED (GLCNAC. N-LINKED (GLCNAC. N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-021-1996 (Rel. 34, Created)
15-JUL-1999 (Rel. 38, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
Arylsulfatase F precursor (EC 3.1.6.-) (ASF).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           591 AA
             MGD; MGI:88075; As1-s.
InterPro; IPR000917; Sulfatase.
Pfam; PF00884; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; PARTIAL.
PROSITE; PS00149; SULFATASE_2; 1.
                                                                                                                                                                  POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DLAGGSTNGTKPLDGFNAWKTISEGHPSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SLAEGQIDEDIQLDGYDIWETISEGLRSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                     Hydrolase; Glycoprotein; Lysosome.

NON_TER 1

NON_CONS 33 34

NOT_SITE 45 45

CARBOHYD 86 86 N LIN

CARBOHYD 177 177 N LIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 370-423 FROM N.A.
                                                                                                                                                                                                                                                                                31727
                                                                                                                                                                                                                                                                                                                                                                        Matches 164; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Genomics 42:192-199(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                             34
45
86
177
189
264
285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens (Human)
                                                                                                                                             33
45
86
177
189
264
285
                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TISSUE-Kidney;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Xp22.3 region.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ARSF_HUMAN
P54793;
                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local
                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                              NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ARSF_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AC NEW TO THE READ OF THE READ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         208 KLSGWVSVPWLLIFSMILFIFLLGYAWFSSH---TSP-----LY------WDCLLMRGH 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             288 INRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNGGQPTA-----GGSNWP 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         73 TSQPHLIFILADDQGFRDVGYHGSE-IKTPTLDKLAAEGVKL-ENYYVQPICTPSRSQFI 130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EITEQPMKAERAGSIMVKEAISFLERHSKETFLLFFSFLHVHTPLPTTDDFTGTSKHGL-311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GYDIWETISEGLRSPRVDILHNIDPIYTKA-----KNGSWAAGYGIWNTAIQSAIRVQHW 454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             131 TGKYQIHTGLQHS----IIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGF--- 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28 TTRPNIVLIMVDDLG1GDLGCYGNDTMRTPHIDRLAREGVRLTQHISAASLCSPSRSAFL 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----TDN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GIYSTQMYTQRVQQILA----SHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRSIIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              341 LRGSKGT-YWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BY SIMILARITY.

"LINKED (GLCNAC. . .) (POTENTIAL)

"LINKED (GLCNAC. . .) (POTENTIAL)

"LINKED (GLCNAC. . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             151;
Franco B., Meroni G., Parenti G., Levilliers J., Bernard L., Gebbia M., Cox L., Maroteaux P., Sheffield L., Rappold G.A., Andria G., Petit C., Ballabio A.;
Andria G., Petit C., Ballabio A.;
A cluster of sulfatese genes on X922.3: mutations in chondrodysplasia punctata (CDPX) and implications for warfarin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 1; Length 591;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2-AMINO-3-OXOPROPIONIC ACID (BY SIMILARITY).
                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1. LINKED (GLCNAC. . .) (P. 858489BF3803A771 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             14.3%; Score 430.5; DB 1; 25.1%; Pred. No. 2.5e-26; ive 86; Mismatches 189;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ARYLSULFATASE F.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              182 YRKEC-MPTRRGFDTFFGS------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR000917; Sulfatase.
Pari; Pr00884; Sulfatase; 1.
PROSITE: PS00523; SULFATASE 1; 1
PROSITE; PS00149; SULFATASE 2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Glycoprotein; Signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; X97868; CAA66462.1; -.
HSSP; P15848; 1FSU.
MIM; 300003; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        66003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best_Local Similarity 25.1
Matches 143; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19
591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         140
51
118
337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      140
51
118
337
591 AA;
```

20;

g ò q ò q

```
TIŠSUE=Placenta;
MEDLINE=92095973; PubMed=1755850;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Chordata;
Primates;
                                                                                                                                                                535 AA; 59814 MW;
                               Hydrolase; Complete proteome
                                                                                                                                                                                                                                                                                                                                        Matches 144; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens (Human).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa;
Mammalia; Eutheria;
                                                                                                                                                                                                                                                                                                        Similarity
PROSITE; PS00149;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (Chondroitinase)
                                                                                                 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GA6S_HUMAN
P34059;
                                                              INIT_MET
ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                        Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GALNS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GA6S_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ID DOT DOT SERVING ON 
      KW
FT
FT
SO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -:- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.
-:- SUBUNIT: MONOMER.
-:- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
-:- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -1- FUNCTION: SHOWS MAXIMAL ACTIVITY AT 57 DEGREES CELSIUS AND PH 8.9. INCUBATION THAT OVERSTEPS 20 MINUTES ABOVE 50 DEGREES CELSIUS LEADS TO ENZYME INACTIVATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
STRAIN-ATCC 15692 / PAO1;
STRAIN-ATCC 15692 / PAO1;
SCHOOLINE-20437337; PubMed-10984043;
Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P., Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M., Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y., Brody L.L., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Raizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.; "Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen."
Nature 406:959-964(2000).
                                                              455 KLLTGNPGYSDWVPPQS----FSNLGPNRWHNERITSSTGKSVWLFNITADPYERVDLSN 510
                                                                                                               470 KAHYVTPVFQ---PPASGGCYVTSL--CRCFGEQVTYH--NPPLLFDLSRDPSESTPLTP 522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
STRAIN—ATCC 15692 / PAOJ;
MEDLINE—95262702; PubMed=7744061;
Bell S., Kehrli H., James P., Staudenmann W., Cook A.M., Leisinger T Kertesz M.A.;
"Purification and characterization of the arylsulfatase synthesized by Pseudomonas aeruginosa PAO during growth in sulfate-free medium and cloning of the arylsulfatase gene (atsA).";
Eur. J. Biochem. 229:385-394(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Arylsulfatase (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Kertesz M.A.;
Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
GRDLMPLLQGNVRHSEHEFLFHYCGSYLHAVRWIPKDDSGSV-
                                                                                                                                                                                                                                                                                                                                                                                                                                         535 AA.
                                                                                                                                                                                                   511 R----YPGIVKKLLRRLSQFNKTAVPVRY 535
                                                                                                                                                                                                                                         Interpro; IPR000917; Sulfatase.
Pfam; PF00884; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; Z48540; CAA88421.2; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AE004456; AAG03573.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pseudomonas aeruginosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HSSP; P15848; 1FSU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (Rel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-287;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATSA OR PA0183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                         ARS_PSEAE
P51691;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     [2]
REVISIONS
                                                                                                                                                                                                                                                                                                                                                                                                      ARS_PSEAE
```

```
225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      226 RYDAGPEALRQERLARLKELGLVEADVEAHPVLALTREWEALEDEERAKSARAMEVYAAM 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----GSNW-----PLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVH 376
                                                                                                                                                                                                                                                          73 TSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITG 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        434 WAAGYGIWNTAIQSAIRVQHWKLLTGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGKSV 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             133 KYQIHTGLQHSII-------RPTQPNCLPLDNATLPQKLKEVGYSTHMVGKW 177
                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HLGFYRKECMPTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCGYD-----LYEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DNAAWD-YDNGIYSTQMYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DERYLDTLPEGFYSSDAFGDKLLQYLKERDQSRPFFAYLPFSAPHWPLQAPREIVEKYRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 346 LENIGRANSYVWYGPRWAQAATAPSRLYKAFTTQGGIRVPALVRYPRLSRQGAISHAFAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           377 ITDWYPTLISLAEGQIDEDIQLDGYDIWETISE---GLRSPRVDILHNIDPIYTKAKNGS
                                                                                                                                                                                                                                                                                                                  1 SKRPNFLVIVADDLGFSDIGAFGGEIATPNLDALAIAGLRLTDFHTASTCSPTRSMLL--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ----RRYAAM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LSCLDEAI NNVTLALKTYGFYNNSIII I YSSDNGGO-----PTAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Craniata; Vertebrata; Euteleostomi;
Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-FEB-1994 (Rel. 28, Created)
01-FEB-1994 (Rel. 28, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
N-acctylgalactosamine-6-sulfatase precursor (EC 3.1.6.4) (N-acctylgalactosamine-6-sulfate sulfatase) (Galactose-6-sulfate sulfatase) (GalNAc6S sulfatase) (Chondroltinsulfatase)
POTENTIAL.
S -> D (IN REF. 1; AA SEQUENCE).
7404F3749E14EA77 CRC64;
                                                                                                                                             Length 535;
                                                                                                                                          ; Score 429; DB 1; Length 535; Pred. No. 2.9e-26; 71; Mismatches 190; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----IININR-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         494 WLFNITADPYERVDLSNRYPGIVKKLLRRLSQF-NKTAV 531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
                                                                                                                                          14.2%; Score 429; 24.9%; Pred. No. 2
```

```
Ogawa T., Tomatsu S., Fukuda S., Yamagishi A., Maruf Rezvi G., Sukegawa K., Kondo N., Suzuki Y., Shimozawa N., Orii T.; "Mucopolysaccharidosis IVA: screening and identification of mutations of the N-acetylgalactosamine-6-sulfate sulfatase gene."; Hum. MOI. Genet. 4:341-349(1995).
Tomatsu S., Fukuda S., Masue M., Sukegawa K., Fukao T., Yamagishi A., Hori T., Iwata H., Ogawa T., Nakashima Y., Hanyu Y., Hashimoto T., Titani K., Oyama R., Suzuki M., Yagi K., Hayashi Y., Orii T., Morquio disease: isolation, characterization and expression of full-length CDNA for human N-acetylgalactosamine-6-sulfate sulfatase."; Biochem. Biophys. Res. Commun. 181:677-683(1991).
                                                                                                                                                                               MEDIINE-95095267; PubMed-8001980; Morris C.P., Guo X.H., Apostolou S., Hopwood J.J., Scott H.S.; Morris C.P., Guo X.H., Apostolou S., Hopwood J.J., Scott H.S.; Morquio A syndrome: cloning, sequence, and structure of the human N-acetylgalactosamine 6-sulfatase (GALNS) gene."; Genomics 22:652-654(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VARIANTS MPS-IVA ARG-77; TRP-90; VAL-96; LEU-151; GLY-230 AND THR-291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Yamada N., Isogai K., Kato Z., Sikegawa K., Kondo N., Suzuki Y., Shimozawa N., Orii T.; "Two new mutations, Q473x and N4875, in a Caucasian patient with mucopolysaccharidosis IVA (Morquio disease)";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Tomateu S., Fukuda S., Cooper A., Wraith J.E., Maruf Rezvi G., Yamadishi A., Yamada N., Kato Z., Isogai K., Sukegawa K., Kondo N., Suzuki Y., Shimozawa N., Oril T.; "Mucopolysaccharidosis IVA: Identification of a common missense mutation Ill3F in the N-Acetylgalactosamine-6-sulfate sulfatase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-95359983; PubMed-7633425; Tomatsu S., Eukuda S., Cooper A., Wraith J.E., Maruf Rezvi G., Yamaqishi A., Yamada N., Kato Z., Isogai K., Sukegawa K., Kondo N., Suzuki Y., Shimozawa N., Orii T., "Mucopolysaccharidosis type IVA: identification of six novel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Kato Z., Yamada N., Isogai K., Sukegawa K., Kondo N., Suzuki Y., Shimozawa N., Orii T.; Mucopolysaccharidosis IVA: four new exonic mutations in patients with N-acetylgalactosamine-6-sulfate sulfatase deficiency."; Am. J. Hum. Genet. 58:950-962(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VARIANTS MPS-IVA CYS-94 AND VAL-97.
MEDIINE=96423834; PubMed=8825435;
Cole D.E.C., Fukuda S., Gordon B.A., Rip J.W., Lecouteur A.N.,
Rupar C.A., Tomatsu S., Ogawa T., Sukegawa K., Orii T.;
"Heteroallelic missense mutations of the galactosamine-6-sulfate
                                                                                                                                                                                                                                                                                                                                            MEDLINE-92395122; PubMed-1522213; Kudud 8., Tomateu H., Ogawa T., Fukuda S., Tomateu S., Masue M., Sukegawa K., Iwata H., Ogawa T., Furi T., Yamagishi A., Hanyu Y., Morooka K., Kiman HashImoto T., Orli T.,
                                                                                                                                                                                                                                                                                                                                                                                                                                    "Mucopolysaccharidosis type IVA. N-acetylgalactosamine-6-sulfate sulfatase exonic point mutations in classical Morquio and mild
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pomatsu S., Fukuda S., Yamagishi A., Cooper A., Wraith J.E.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARIANTS MPS-IVA ALA-138; SER-151 AND LEU-151.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      mutations among non-Japanese patients.";
Hum. Mol. Genet. 4:741-743(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Clin. Invest. 90:1049-1053(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         J. Hum. Genet. 57:556-563(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VARIANTS MPS-IVA.
MEDLINE-95397840; PubMed=7668283;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-96047158; PubMed-7581409;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-95315929; PubMed-7795586;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-96216146; PubMed-8651279;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   fum. Mutat. 6:195-196(1995).
                                                                                                                                                                                                                                                                                                                         VARIANT MPS-IVA LYS-204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARIANT MPS-IVA SER-487
                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VARIANTS MPS-IVA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gene.";
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Fifteen polymorphisms in the N-acetylgalactosamine-6-sulfate
"Fifteen polymorphisms in the N-acetylgalactosamine-6-sulfate
"Tifteen polymorphisms in the N-acetylgalactosamine diagnostic implications in Morquio disease.";
"Hum. Mutat. Suppl. 1:542-546(1998).

-!- ATALYTIC ACTIVITY: Hydrolysis of the 6-sulfate groups of the N-acetyl-D-galactose 6-sulfate units of chondroitin sulfate and of the D-galactose 6-sulfate units of keratan sulfate and of the D-galactose 6-sulfate units of keratan sulfate and of the D-galactose 6-sulfate units of keratan sulfate."

-!- SUBGUIT: OLIGOMER OF DISULFIDE LINKED 40- AND 15-kDa POLYPEPTIDES.
-!- SUBGELGUAR LOCATION: Lysosomal.
-!- SUBGELGUAR LOCATION: Lysosomal.
-!- Type ITA (MPS-IVA) (also known as Morquio A syndrome) which is characterized by specific spondyloepiphyseal dysplasia, short trunk dwarfism, coxa valga, odontoid hypoplasia, corneal opacities, preservation of intelligence, and excessive urinary excretion of keratan sulfate and chondroitin-6-sulfate. Severely affected patients usually die of cardiopulmonary disturbance or cervical cord compression in the second or third decade of life.
                                                                                                                                                                    "Identification of 31 novel mutations in the N-acetylgalactosamine-6-sulfatase gene reveals excessive allelic heterogeneity among patients with Morquio A syndrome.";
Hum. Mutat. 10:223-232(1997).
sulfatase (GALNS) gene in a mild form of Morquio disease (MPS IVA)."; Am. J. Med. Genet. 63:558-565(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Yamada N., Fukuda S., Tomatsu S., Muller V., Hopwood J.J., Nelson J., Kato Z., Yamadishi A., Sukegawa K., Kondo N., Orii T.; "Molecular heterogeneity in mucopolysaccharidosis IVA in Australia and Northern Ireland: nine novel mutations including T312S, a common allele that confers a mild phenotype.";
                                                          VARIANTS MPS-IVA.
MEDLINE=97442274; PubMed=9298823;
Bunge S., Kleijer W.J., Tylki-Szymanska A., Steglich C., Beck M.,
Tomatsu S., Fukuda S., Poorthuis B.J.H.M., Czartoryska B., Orii T.,
                                                                                                                                                                                                                                                                                                                      Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Ferreira P.,
di Natale P., Tortora P., Fujimoto A., Kato Z., Yamada N., Isogai
Amangishi A., Sukegawa K., Suzuki Y., Shimozawa N., Kondo N.,
Sly W.S., Orii T.,
"Fourteen novel mucopolysaccharidosis IVA producing mutations in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=98112415; Pubmed=9452036;
Tomatsu S., Fukuda S., Cooper A., Wraith J.E., Yamagishi A., Kato
Yamada N., Isogai K., Sukegawa K., Suzuki Y., Shimozawa N., Kondo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VARIANTS MPS-IVA.
MEDLINE=98180718; PubMed=9521421;
                                                                                                                                                                                                                                                                                                      MEDLINE=98041700; PubMed=9375852;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BAA04535.1; JOINED.
BAA04535.1; JOINED.
BAA04535.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BAA04535.1; JOINED.
BAA04535.1; JOINED.
BAA04535.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BAA04535.1; JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BAA04535.1; JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                                  GALNS gene.";
Hum. Mutat. 10:368-375(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hum. Mutat. 11:202-208(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARIANTS SER-393 AND MET-488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            D17629; BAA04535.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BAA04535.1;
                                                                                                                                                                                                                                                                                  VARIANTS MPS-IVA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                D17616;
D17617;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    D17623;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             D17618;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               D17619;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       D17621;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL;
```

œ

```
423 NVSGV---TTHN---LEDHTKLPLIFHLGRDPGERFPLSFASAEYQEALSRITSVVQQHQ 476
 SFSNLGPNRWHNERITSSTGKSVWLFNITADPYERVDL---SNRYPGIVKKLLRRLSQFN 527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        D55B627983A4C4D6 CRC64;
                                                                                                                    567 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Interpro: IPR000917; Sulfatase.
Pfam: PF00884; Sulfatase: 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
MEDLINE-89357267; PubMed=2767335;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Glycoprotein.
                                           547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .;
₩
                                                         : | | : | 1 | EALVP-----AOPOLN 487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; M28404; AAA30036.1; -.
                                           KTAVPVRYPPKDPRSNPRLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        62477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              567
115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             173
179
228
542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            173
179
228
542
567 AA;
                                                                                                                                                                                                                             Strongylocentrotus.
NCBI_TaxID=7668;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Matches 120; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20
115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HSSP; P15289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hydrolase;
                                                                                                                    ARS_STRPU
P50473:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIGNAL
                                                                                                                                                                                    (ARS).
                                          528
                                                               477
                                                                                                         ARS_STRPU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28
                    g
                                           ò
                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                       20;
                                                                                                                                                                                                                                                                             SIMILARITY).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
G-YR (IN MPS-IVA; SEVERE FORM).
/FIId=VAR_007172.
D-> N (IN MPS-IVA; MILD FORM).
/FIId=VAR_007173.
L-> M (IN MPS-IVA; ASSOCIATED WITH S-409
IN A PATIENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DDQGFRDVGYHGSEIK-TPTLDKLAAEGVKLENYY-VQPICTPSKSQFITGKYQIHTGL- 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           : | : :| : | : | TINAHARNAYIPQEIVGGIPDSEQLLPELLKKAGYVSKIVGKWHLG-HRPQFHPLKHGFD 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      295 AMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNG----GQPTAGGSNWPLRGSKGTYWE 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEG 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LRSPRVDILHNIDPIYTKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLIGNPGYSDWVPPQ 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----QHSIIRPTQ-PNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFD 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ----TQMYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRSIININRRRYA 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RIMDRPIFYYRGDTLMA-ATLGQHKAHFWTWTNS------WENFRQG---IDFCPGQ 422
                                                                                                                                                                                                                                                                                                                                                                                                                                            24 MLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLAQAGEKLEPSTTSTSQPHLIFILA 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 MAAVVAATRWWQLLL-------VLSAAG---MGASGAPQPPNILLLLM 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam: PF00884; Sulfatass: 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
Hydrolase; Glycoprotein; Signal; Lysosome; Mucopolysaccharidosis; Disease mutation; Polymorphism:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               202 EANLTQIYLQEALDFIKRQARHHPFFLYWAVDATHAPVYASKPF-----LGTSQRGRYG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          195 TFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNAA------WDYDNGIYS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GGMREPALAWWPGHVTAGOVSHOLGSIMDLFTTSLALAGLTPPSDRAIDGLNLLPTLLQG
                                                                                                                                                                                                                                                                                                                                                                                                                       Indels 109;
                                                                                                                                                                                                                                                           N-ACETYLGALACTOSAMINE-6-SULFATASE.
2-AMINO-3-OXOPROPIONIC ACID (BY
                                                                                                                                                                                                                                                                                                                                                                                                 Score 422.5; DB 1; Length 522;
Pred. No. 9e-26;
; Mismatches 225; Indels 109
                                                                                                                                                                                                                                                                                                                                                                           /FTId-VAR_007174.
                                                   JOINED.
                                         JOINED
                                                                                                                                                                                                                                                                                                                                                                                                                      84;
                                                                                                                                                                                 InterPro; IPR000917; Sulfatase.
                                                                                                                                                                                                                                                                                                                                                                                                14.0%;
25.4%;
                                                  EMBL, U06079, AACS1350.11, 37
EMBL, U06080, AACS1350.11, 37
EMBL, U06081, AACS1350.11, 37
EMBL, U06083, AACS1350.11, 37
EMBL, U06084, AACS1350.11, 37
EMBL, U06084, AACS1350.11, 37
                              U06088; AAC51350.1;
U06078; AAC51350.1;
        BAA04535.1;
BAA04535.1;
                                                                                                                            EMBL; U06086; AAC51350.1;
EMBL; U06087; AAC51350.1;
PIR; JQ1299; JQ1299.
HSSP; P15848; 1FSU.
                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                          522
                                                                                                                                                                                                                                                                                         204
                                                                                                                                                                                                                                                                                                              47
                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                                                        204
423
47
                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                       67
                                                                                                                                                                                                                                                                                                                                                                                                                      Matches 142;
                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                    MOD_RES
                                                                                                                                                                                                                                                                                                                                    VARIANT
                                                                                                                                                                                                                                                                                                                                                         VARIANT
                                                                                                                                                                                                                                                  SIGNAL
                                          EMBL;
                                EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    376
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LLAQAGEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKTP-TLDKLAAEGVKL-EN 115
                                                                                                                                                                            Strongylocentrotus purpuratus (Purple sea urchin).
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoldea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
01-ocT-1996 (Rel. 34, Created)
01-ocT-1996 (Rel. 34, Last sequence update)
01-oCT-1996 (Rel. 34, Last annotation update)
Arylsulfatase precursor (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CELLS AND THEIR PRECURSORS.

DEVELOPMENTAL STAGE: LOW LEVELS ARE FOUND AT MESENCHYME BLASTU STAGE (24 HR), LEVELS INCREASE BY LATE GASTRULA STAGE AND ARE MAINTAINED AT PLUTEUS STAGE.

SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Yang Q., Angerer L.M., Angerer R.C.;
"Structure and tissue-specific developmental expression of a sea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC. . .) (POTENTIAL)
N-LINKED (GLCNAC. . .) (POTENTIAL)
N-LINKED (GLCNAC. . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13.8%; Score 416.5; DB 1; Length 567; 29.7%; Pred. No. 3e-25; tive 66; Mismatches 175; Indels 43;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ARYLSULFATASE.
2-AMINO-3-OXOPROPIONIC ACID (BY
```

```
Biochim. Biophys. Acta 997:199-205(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Oinuma T.,
                                                                                                                                                                    SDDFSCTSRRGRYGDNLREMDQAIEQIVTTLVDNDIDDNTVIFFTSDHGPHREYCGE--- 333
                                                                                                                                                                                                        01-NOV-1988 (Rel. 09, Created)
01-APR-1990 (Rel. 14, Last sequence update)
01-OCT-1996 (Rel. 34, Last annotation update)
Steryl-sulfatase precursor (EC 31.6.2) (Steroid sulfatase) (Steryl-sulfate sulfohydrolase) (Arylsulfatase C) (ASC).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Cloning and expression of steroid sulfatase cDNA and the frequent occurrence of deletions in STS deficiency: implications for X-Y
 YYVQPICTPSRSQFITGKYQIHTGL--QHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHM
                                                            AWDYDNGIYSTQMYTQR-VQQIL-----ASHNPTKPIFLYIAYQAVHSPLQAPGRYFE
                                                                                                                                            281 HYRSIININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNG-----GQPTA
                                                                                                                                                                                          335 GGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDE
                       YSGDSVCTPSRSAIVTGRQPIRTGVYGEERIFLPWTTTGLPLYEVTIAEAMKGAGYTTGM
                                              VGKWHLGFYRKEC-----MPTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNA
                                                                                                                     223 CFLYYNSTSVAQPFQHKGLTQLLRDDTVGFIEDNVNKPFFMYVSFAHMHTSL-----FS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               m
m
                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Yen P.H., Allen E., Marsh B., Mohandas T., Wang N., Taggart R.T., Shapiro L.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-89077541; PubMed-3203382;
Yen P.H., Marsh B., Allen E., Tsai S.P., Ellison J., Connolly L.,
Meiswanger K., Shapiro L.J.;
"The human X-linked steroid sulfatase gene and a Y-encoded
pseudogene: evidence for an inversion of the Y chromosome during
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Cloning and expression of human steroid-sulfatase. Membrane topology, glycosylation, and subcellular distribution in BHK-21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDLINE-89340479; PubMed=2668275;
Stein C., Hille A., Seidel J., Rijnbout S., Waheed A., Schmidt
Geuze H., von Figura K.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-89352671; PubMed-2765556; Kawano J.-I., Kotani T., Ohtaki S., Minamino N., Matsuo H.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Characterization of rat and human steroid sulfatases.";
                                                                                                                                                                                                                                                                  393 DRIYDGKCLKSVLLEGASSPHDDFFYYCKDTLMAVRVGKYKAHF 436
                                                                                                                                                                                                                                          395 DIQLDGYDIWETISEGLRSPRVDILHNIDPIYTKAKNGSWAAGY 438
                                                                                                                                                                                                                                                                                                                             583 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 134-174 AND 461-583 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        J. Biol. Chem. 264:13865-13872(1989).
                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=87187642; PubMed=3032454;
                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           cell 55:1123-1135(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ceil 49:443-454(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  [2]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE OF 22-45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              primate evolution,
                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 rissue=Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      interchange
                                                                                                                                                                                                                                                                                                                           STS_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Aikawa
 116
                                                                                              529
                                                                                                                                                                    277
                                                                                                                                                                                                                                                                                                                 STS_HUMAN
                                                                                                                                                                                                                                                                                                                              ò
                       පු
                                               ò
                                                                   셤
                                                                                            ò
                                                                                                                   g
                                                                                                                                            ò
                                                                                                                                                                 a
                                                                                                                                                                                          ò
                                                                                                                                                                                                                  염
                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                  g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hydrolase; Microsome; Transmembrane; Glycoprotein; Steroid metabolism; Pregnancy; Signal; Disease mutation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     125 SRSQFITGKYQIHTGLQH----SIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHL 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EPSTTSTSQPHLIFILADDQGFRDVGYHGSE-IKTPTLDKLAAEGVKL-ENYYVQPICTP 124
                                                                                                                                                                                                H(2)O = 3-beta-hydroxyandrost'5-en-17-one + sulfate.
SUBUNIT: HOWODIMER.
SUBUNIT: HOWODIMER.
SUBURIT: HOWODIMER.
SUBURIT: HOWODIMER.
SEVERAL MEMBRANE-SPANION MICROSOMAL MEMBRANE. THE SEQUENCE SHOWS
BYOTEIN IN THE MICROSOMAL MEMBRANE.
PROTEIN IN THE MICROSOMAL MEMBRANE.
SUBSARS: DEFECTS IN STS ARE A CAUGE OF A VISIBLE PHENOTYPE OF SCALY SKIN, X-LINKED ICHTHYOSIS (XLI).
SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | : : |:|::| ::|| | | | : |::|| EAESHAASRPNIILVMADDLGIGDPGCYGNKTIRTPNIDRLASGGVKLTQHLAASPLCTP 77
                                                                                                                                       ESTROGENS
                                      Basler E., Grompe M., Parenti G., Yates J., Ballabio A.; "Identification of point mutations in the steroid sulfatase gene of three patients with X-linked ichthyosis."; Am. J. Hum. Genet. 50:483-491(1992).
                                                                                                                                                                             CATALYTIC ACTIVITY: 3-beta-hydroxyandrost-5-en-17-one 3-sulfate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2-AMINO-3-OXOPROPIONIC ACID (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13.7%; Score 413; DB 1; Length 583; 25.1%; Pred. No. 5.9e-25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                78; Mismatches 201; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              74746AFA9D21A0A6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
S-D (IN XLI).
FTIG-VAR_007240.
W->R (IN XLI).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STERYL-SULFATASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /FTId=VAR_007241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     007242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C -> X (IN XLI)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A -> E (IN REF.
VARIANTS XLI LEU-341; ARG-372 AND TYR-446.
MEDLINE-92170784; PubMed=1539590;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /FTIG=VAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR000917; Sulfatase.
Para; PF00884; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  65492 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL, J04964; AAA60597.1; --
EMBL, M16505; AAA60596.1; --
EMBL, M23945; AAA60598.1; --
EMBL, M23556; AAA60599.1; --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     583
75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        136
206
234
47
259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        333
459
341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PIR; A32641; A32641.
PIR; A25961; A25961.
PIR; S05415; S05415.
HSSP; P15848; IFSU:
MIM; 308100; -
                                                                                                                                                          DURING PREGNANCY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23
583 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        136
185
213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /ARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /ARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
```

22;

```
336
                                                                                                                                                                                                                                                                                                                                                  419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KAFFFTP-----NFNPVGSNGCFATHVCFCFGSYVTHDDPLLFDISKDPRERNPL 515
                                                                                                                                                                                                                                                                                                                 SNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDI 396
                                                                                                                                                                                                                                                                                                                                                                                   QLDGYDIWETISEGLRSPRVD--ILHNIDPIYTKAKNGSWAAGYGIWNTAIQSAIRVQHW 454
                                                                                                                                                                                                                                                                                                                                                                                                                464
Q---PMSYDN----LTQRLTVEAAQFI-QRNTETPFLLVLSYLHVHTALFSSKDFAGKSQH 305
                                                                                                                                                                                                                                                                                                                                       GM---SCHSKTDFCHHPLHHGFNYFYGISLTNLRDCKPGEGSVFTTGFKRLVFLPLQIVG
                                                                                                                -----HX----KCDSPGMCGYDLYE
                                                                                                                                               195 VTLLTLAALNCLGLLHVPLGVFFSLLFLAALILTLFLGFLHYFRPLNCFM--MRNYEIIQ
                                                                                                                                                                                NDNAAWDYDNGIYSTQMYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRS
                                                                                                                                                                                                                                                  285 IININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNGGQ------PTAGG
                                                                                                                                                                                                                                                                    GV-----YGDAVEEMDWSVGQ1LNLLDELRLANDTLIYFTSDQGAHVEEVSSKGEIHGG
                                                                                                                                                                                                                                                                                                                                                                                                                  120 IIDGRDLMPLL-EG-KSQRSDHEFLFHYCNAYLNAVR------WHPQNSTSI----W
                                                                                                                                                                                                                                                                                                                                                                                                                                                   KLLTGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGKSV-----WLFNITADPYERVDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Chordata; Craniata; Vertebrata; Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-98072521; PubMed-9409863;
Parenti G., Buttitta P., Meroni G., Franco B., Bernard L.,
Rizzolo M.G., Brunetti-Pierri N., Ballabio A., Andria G.;
"X-linked recessive chondrodysplasia punctata due to a new point
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Franco B., Meroni G., Parenti G., Levilliers J., Bernard L., Gebbia M., Cox L., Maroteaux P., Sheffield L., Rappold G.A., Andria G., Petit C., Ballabio A.;
Andria G., Petit C., Ballabio A.;
A cluster of sulfatase genes on Xp22.3: mutations in chondrodysplasia punctata (CDPX) and implications for warfarin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Daniele A., Parenti G., D'Addio M., Andria G., Ballabio A.,
Meroni G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Catarrhini; Hominidae; Homo.
                                                -----MPTRRGFDTFFGSLL-----GSGDYYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-07T-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
16-0CT-2001 (Rel. 40, Last annotation update)
Arylsulfatase E precursor (EC 3.1.6.-) (ASE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    589 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A., AND VARIANTS CDPX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          the ARSE gene.";
Genet. 73:139-143(1997)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S----NRYPGIVKKLLRRLSQFNKT 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TPASEPRFYEILKVMQEAADRHTQT 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE-Kidney;
MEDLINE-95236447; PubMed-7720070;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHARACTERIZATION.
MEDLINE-98163440; Pubmed-9497243;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARIANT CDPX TYR-492.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Cell 81:15-25(1995).
                                                GFYRKEC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ARSE_HUMAN
P51690;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           embryopat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            mutation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ARSE_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      516
                 78
                                                                               138
                                                                                                                508
                                                                                                                                                                                  225
                                                                                                                                                                                                                 253
                                                                                                                                                                                                                                                                                   306
                                                                                                                                                                                                                                                                                                                  337
                                                                                                                                                                                                                                                                                                                                                  360
                                                                                                                                                                                                                                                                                                                                                                                   397
                                                                                                                                                                                                                                                                                                                                                                                                                                                    455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   509
                                                180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                               g
                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                ò
                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                 à
                                                ð
                                                                                                                  ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                     SUBCELLULAR LOCATION: GOLGI STACK.

SUBCELLULAR LOCATION: GOLGI STACK.

PITSSUE SPECIFICITY: EXPRESSED IN THE PANCREAS, LIVER, AND KIDNEY.

PTHIS SPECIFICITY: EXPRESSED IN THE PANCREAS.

DISEASE: DEFECTS IN ARSE ARE THE CAUSE OF X-LINKED RECESSIVE

ENDNDROPSPELASIA PUNCTATA (CDPX). A COMENITAL DEFECT OF BONE AND

CARTILAGE DEVELOPMENT CHARACTERIZED BY ABERRANT BONE

MINERALIZATION, SEVERE UNDERDEVELOPMENT OF NASAL CARTILAGE, AND

BISTALL PHALANCEAL HYPOPLASIA. THIS DISEASE CAN ALSO BE INDUCED BY

INHIBITION WITH THE DRUG WARRARIN.

SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LEPSTT---STSQPHLIFILADDQGFRDVGYHGSE-IKTPTLDKLAAEGVKL-ENYYVQP 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121 ICTPSRSQFITGKYQIHTGLQHSI-IRPTQ----PNCLPLDNATLPQKLKEVGYSTHMVG 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            84
                                                                                PH
                                                    FUNCTION: MAY BE ESSENTIAL FOR THE CORRECT COMPOSITION OF CARTILLAGE AND BONE MATRY DURING BEVELOMENT. HAS NO ACTIVITY TOWARD STEROLD SULFATES AND SHOWS MAXIMAL ACTIVITY AT NEUTRAL PHIS ALMOST COMPLETELY INACTIVATED AFTER 10 MINUTES AT 50 DEGREES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                             ENZYME REGULATION: INHIBITED BY MILLIMOLAR CONCENTRATIONS OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
"Biochemical characterization of arylsulfatase E and functional
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      161;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 589;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACID (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19F7D8DA999E528F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2-AMINO-3-OXOPROPIONIC AC
SMILARTY).
N-LINKED (GLCNAC...) (!
             in patients with X-linked
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disease mutation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     188;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13.5%; Score 407.5; DB 1
24.0%; Pred. No. 1.6e-24;
ive 92; Mismatches 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /FTId=VAR_007310.
G -> R (IN CDPX).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -> V (IN CDPX)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /FTId=VAR_007308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /FTId=VAR_007311
C -> Y (IN CDPX)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P (IN CDPX)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -> R (IN CDPX)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      007307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL.
ARYLSULFATASE E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FTIG-VAR_0073
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /FTId=VAR_
R -> P (IN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FTIG=VAR
G -> V (IN
                                         J. Hum. Genet. 62:562-577(1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Incerty, Pro0884; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
Hvdrolase; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR000917; Sulfatase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                W.
                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; X83573; CAA58556.1; -. HSSP; P15848; 1FSU.
             analysis of mutations found
chondrodysplasia punctata.";
Am. J. Hum. Genet. 621:562-57'
-!- FUNCTION: MAY BE ESSENTI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                65683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       589
147
86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               58
125
258
344
12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                589 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
nes 139; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               58
125
258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     111
                                                                                                                                                                                                                                                                                                                                                                                                                                                           MIM; 300180;
MIM; 302950;
MIM; 302940;
                                                                                                                                            WARFARIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ô
```

25;

11

```
.
H
                                                                                                                                                                                                                                                                                                                                                        359 ENQLGNTQYGGWNGIYKGGKGMGGWEGGIRVPGIFRWPGVLPAGRVIGEPTSLMDVFPTV 418
                                                                                                                                                                                                                                                                                                                                                                                                                       384
                                                                                                                                                                                                                                                                                                                                                                                              385 ISLAEGOIDEDIQLDGYDIWE-TISEGLRSPRVDILHNIDPIYTKAKNGSWAAGYGIWNT 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 444 AIQSAIRVQHWKLLTGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGKSV-----WLFN 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RDRGTM----WKV------HFVTP-VFQPEGAGACYGRKVCPCFGEKVVHHDPPLLFD 515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEVERAL MEMBRANE-SPANNING DOMAINS THAT COULD SERVE TO ANCHOR THE PROTEIN IN THE MICROSOMAL MEMBRANE.
SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Rattus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sulfatase) (Steryl.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CATALYTIC ACTIVITY: 3-beta-hydroxyandrost-5-en-17-one 3-sulfate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Oinuma
PGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQA
                                                                                                                                                                                                                       P-MC----FQR------TTPLILQEVASFL-KRNKHGPFLLFVSFLHVHIPLIT
                                                                                                                                                                                                                                                                                                                              ----GQPTAGGSNWPLRGSKGT-YWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTL
                                                                                                                       -----DSYTHYKC-----DS
                                                                                                                                                                                                                                                                                        ----MENFLGKSLHGL----YGDNVEEMDWMVGRILDTLDVEGLSNSTLIYFTSDHGGSL
                                                                                                                                                     205 ALVALTLVAGKLTHLIPVSWMPVIWSALSAVLLLASSYFVGALIVHADCFLMRNHTITEQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Ratt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          THE SEQUENCE SHOWS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .
H
                                                                                                                                                                                                                                                        PGRYFEHY -- RSIININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Li X.M., Salido E.C., Gong Y., Yen P.H., Kitada Y., Serikawa
Shapiro L.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ή.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Characterization of rat and human steroid sulfatases.";
Blochim. Blophys. Acta 997:199-205(1989).
-!- FUNCTION: CONVERSION OF SULFATED STEROID PRECURSORS TO DURING PREGNANCY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-89352671; Pubmed-2765556;
Kawano J.-I., Kotani T., Ohtaki S., Minamino N., Matsuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        H(2)0 = 3-beta-hydroxyandrost-5-en-17-one + sulfate.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-APR-1990 (Rel. 14, Created)
01-CT-1996 (Rel. 34, Last sequence update)
01-CCT-1996 (Rel. 34, Last annotation update)
Steryl-sulfatase precursor (EC 3.1.6.2) (Steroid sulfate sulfohydrolase) (Arylsulfatase C) (ASC).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   498 ITADPYERVDLSNRYPGIVKKLLRRLSQ----FNKTAVPV 533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 516 LSRDPSETHILTPASEPVFYOVMERVOOAVWEHORTLSPV 555
                                                   KWHLGFYRKECM----PTRRGFDTFFG---SLLGS----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUBUNIT: HOMODIMER.
SUBCELLULAR LOCATION: MICROSOMAL MEMBRANE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        577 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-SPRAGUE-DAWLEY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE OF 20-45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TISSUE=Liver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STS_RAT
P15589;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Aikawa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Π
                                                                                 145
                                                                                                                                                                                                                        265
                                                                                                                                                                                                                                                          275
                                                                                                                   204
                                                                                                                                                                                        215
                                                                                                                                                                                                                                                                                          307
                                                                                                                                                                                                                                                                                                                            330
                 윤
                                                   ö
                                                                                 8
                                                                                                                   õ
                                                                                                                                                   셤
                                                                                                                                                                                       ò
                                                                                                                                                                                                                   Q
                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                          å
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20;
                                                                                                                                                                                                                                       Hydrolase; Microsome; Transmembrane; Glycoprotein; Steroid metabolism;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             143 ASDFCHHPGRHGFDRFLGTPT-----TNLRDCKPGGGTVFGSAQQVFVVLPMNILGAVL 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            371 WEGGIRVPGLVRWPGVIVPGQEVEEPTSNMDVFPTVARLAGAELPTDRVIDGRDLMPLLL 430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PHLIFILADDQGFRDVGYHGSE-IKTPTLDKLAAEGVKL-ENYYVQPICTPSRSQFITGK 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 134 YQIHTGL-QHS----IIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECM- 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ---MCGYDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----TOR----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               197 LAMALARWAGLARPPGWVFGVTVAAMAAVGGAYVAFLYHFRPANCFLMADFTITQQPTDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  311 VEEMDWAVGOVLATLDKLGLANNTLVYLTSDHGAHVEELGPNGERHGGSNGIYRGGKANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 349 WEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EGLRSPRVDILHNIDPIYTKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGNPGYSDWVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       431 GHVQHSEHEFLFHYCNAYLSAV-----AWRPHNSSSV----WKAFYFTPNFD---P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --VQQILAS------HNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRSIININRRRYAAM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LSCLDEAINNVTLALKTYGFYNNSIIIYSSDNGGQ----PTA---GGSNWPLRGSKGTY
                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (GLCNAC. .) (POTENTIAL)
N-LINKED (GLCNAC. .) (POTENTIAL)
N-LINKED (GLCNAC. ..) (POTENTIAL)
07562EADEDCOB716 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           66; Mismatches 193; Indels 130;
                                                                                                                                                                                                                                                                                                (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 577;
                                                                                                                                                                                                                                                                                             2-AMINO-3-OXOPROPIONIC ACID SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---WLFNITADPYERVDLS 509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 403; DB 1;
Pred. No. 3.6e-24;
                                                                                                                                                                                                                                                                                 STERYL-SULFATASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     223 YENDNAAW-----DYDNGIYSTQMY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----PTRRGFDTFFGSLLGSGDYYTHYKCDSPG--
                                                                                                                                                                                                                                                                                                                                          POTENTIAL.
POTENTIAL.
                                                                                                                                                                         InterPro; IPR000917; Sulfatase.
Pfam; PF00084; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POSFSNLGPNRWHNERITSSTGKSV ---
                                                                                                                                                                                                                                                                                                                                                                                                                    Ψ.
                                                                                                                                EMBL; U37138; AAC53097.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              13.4%;
26.2%;
                                                                                                                                                                                                                                                                                                                                                                       46
332
458
62678 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             al Similarity 26.29
138; Conservative
                                                                                                                                                                                                                                                                   19
577
74
                                                                                                                                               PIR; S05414; S05414.
                                                                                                                                                                                                                                                                                                                          135
182
213
46
332
458
577 AA;
                                                                                                                                                               HSSP; P15848; 1FSU
                                                                                                                                                                                                                                                       Signal.
                                                                                                                                                                                                                                                       Pregnancy;
                                                                                                                                                                                                                                                                                                                          ACT_SITE
TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                               MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local
                                                                                                                                                                                                                                                                   SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       С
ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

12

RESULT

```
199 QLWGYTQFLALGILTLAAGQTCGFFSVSARAVTGMAGVGCLFFISWYSSFGFVRRWNCIL 258
                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HSSP; P15848; 1FSU.
EcoGene; EG10089; aslA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                         ----XZHNPTK---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-K12;
                                                                                                                                                                                                                                                                                                                                                       ASLA_ECOLI
                                                                                                                                                                                                                                                                                                                                 RESULT 13
ASLA_ECOLI
                                                                                               319
                                                                                                                                                                                                                                             478
                        251
                                                                       281
                                                                                            g
                                                                                                                                            a
                                                                                                                                                                                             g
                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                       δ
                                              8
                                                                       à
                                                                                                                       δ
                                                                                                                                                                      ŏ
                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SRSQFITGKYQIHTGLQHSI-IRPTQPNC----LPLDNATLPQKLKEVGYSTHMVGKWHL 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       180 GFYRKECM-----PTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNA--- 228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GV---NCASRGDHCHHPLNHGFDYFYGM----PFTLTNDCD-PGRPP----EVDAALRA 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----YSTQMYTQRVQQIL 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EPSTISISQPHLIFILADDQGFRDVGYHGSE-IKTPTLDKLAAEGVKL-ENYYVQPICTP 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SRAAFLTGRHSFRSGMDASNGYRALQWNAGSGGLPENETTFARILQQHGYATGLIGKWHQ 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13.2%; Score 399; DB 1; Length 593;
24.4%; Pred. No. 7.7e-24;
ive 87; Mismatches 191; Indels 168; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalla; Eutherla; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                Cell #1:15-25(1995).
-!- SUBCELLULAR LOCATION: Lysosomal (Potential).
-!- TISSUE SPECIFICITY: EXPRESSED IN THE PANCREAS, KIDNEY, LIVER,
LUNG, PLACENTA, BRAIN, AND HEART.
-!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
74CD9227952BE5DA CRC64;
                                                                                                                                                                                                                                            warfarin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2-AMINO-3-OXOPROPIONIC ACID (BY SIMILARITY).
                                                                                                                                                                                          Franco B., Meroni G., Parenti G., Levilliers J., Bernard L. Gebbia M., Cox L., Maroteaux P., Sheffield L., Rappold G.A. Andria G., Petit C., Ballablo A.; A cluster of sulfatase genes on Xp22.3; mutations in chondrodysplasia punctata (CDPX) and implications for warfa.
                                            01-OCT-1996 (Rel. 34, Last sequence update)
01-OCT-1996 (Rel. 34, Last annotation update)
Arylsulfatase D precursor (EC 3.1.6.-) (ASD).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL.
ARYLSULFATASE D.
           593 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Lysosome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
           PRT;
                                                                                                                                                                   TISSUE-Kidney;
MEDLINE-95236447; PubMed-7720070;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF00884; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hydrolase; Signal; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -AWDYDN----GI------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR000917; Sulfatase.
                                   (Rel. 34, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      65071 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; X83572; CAA58555.1; -. HSSP; P15848; 1FSU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               150
89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
128
347
                                                                                               Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
128
347
593 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                 NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  34
150
89
        ARSD_HUMAN
P51689;
                                  01-OCT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Simi
Matches 144;
                                                                                                                                                                                                                                                        embryopathy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIGNAL
ARSD_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      229
            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Murphy H.R., Kalman M., Cashel M.;
Submitted (APR-1992) to the EMBL/GenBank/DDBJ databases.
-!- CATALYTIC ACTIVITY: A phenol sulfate + H(2)O = a phenol + sulfate.
-!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
-!- CAUTION: REF. 2 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 462
ONWARD AND IS SHORTER (475 AA) DUE TO A FRAMESHIFT.
                                                                                                                                                                                                                                                                                                                                                                                                                                       --WKV------HYTTPQ-----FHPEERGLLTAEASAHAEWGGVTHHRPPLLFD 518
--PIFLYIAYQAVHSPLQAPGRYF---E 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    393 DEDIQLDGYDIWETISEG-LRSPRVDILHNIDPIYTKAKNGSWAAGYGIWNTAIQSAIRV 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  430 PQDRVIDGHSLVPLLQGAEARSAHEFLFHYCGQHLHAAR--------WHQKDSGSV-- 477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 452 QHWKLLTGNPGYSDWVPPQSFSNLGPNRWHNER----ITSSTGKSVW-----LFN 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
STRAIN=XIZ - MG1655;
STRAIN=XIZ - MG1655;
MEDLINE-92358234; PubMed-1379743;
Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
"Analysis of the Escherichia coli genome: DNA sequence of the region from 84.5 to 86.5 minutes";
Science 257:771-778(1992).
                                                                                                                                                                                                                                          | ::||:|| H::::||:|| H:::::||:|| H:::::||:|| H:::::||:|| H::::::||:|| H:::::||:|| H:::::||:|| H::::||:|| H:::::||:|| H:::::|| H:::::||:|| H:::::|| H:::::||:|| H:::::||
                                                                                    259 MRNHDVTEQPMVLEKTASLMLKEAVSYIERHKHGPFLLFLSLLHVHIPLVTTSAFLGKSQ
                                                                                                                                                                                   HYRSIININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNGGQPTA----
                                                                                                                                                                                                                                                                                                                                                                           335 -GGSNWPLRGSKGT-YWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4.98 ITADPYE----RVDLSNRYPGIVKKLLRRLSQFNKTAVPVRYPPKDPRSN 543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-MAY-1992 (Rel. 22, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
11-GOT-2001 (Rel. 40, Last annotation update)
Arylsulfatase (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase).
ASLA OR ATSA OR B3801.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       551 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; M87049; AAA67597.1; -.
EMBL; AE000456; AAC76804.1; -.
EMBL; M90498; AAC323036.1; ALT_FRAME.
PIR; S30691; S30691.
PIR; S27555; S27555.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
```

```
282 Y-----RSIININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNGGQ--- 331
            FROM N.A.
           SEQUENCE FROM STRAIN-BALB/C;
                                                                                                                                                                                                                                                                                                                                                                         Pregnancy;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                           ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                  SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                           29;
                                                                                                                                                                                                                                                                                                                                                                           DKPFFLYYGTRCCH-------PDNYPNAKYAGSSPARTSYGDCMVEMNDVFANLYKTL 340
                                                                                                                                                                                                                                                                                                                                                                                                                 ENYYVQPICTPSRSQFITGKYQIHTGLQHSIIRPT---QPNCLPLDNATLPQKLKEVGYS 170
                                                                                                                                                                                                                                             THMVGKWHLGFYRKECMPTRRGFDTF--FGSLLGSGDYYTHYKCDSPGMCGYDLYENDNA 228
                                                                                                                                                                                                                                                                                            183 TOAIGKWHMG-ENKESOPONVGFDDFRGFNSV---SDMYTEWR------DVHVNPEV 229
                                                                                                                                                                                                                                                                                                                                                                                                                                              TVCKEP-----VHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNI 422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              489 QGGFTGTVMQTAGSSV--FNLYTDPQESDSIGVRHIPMGVPLQTEMHAYME--ILKKYPP 544
                                                                                                                                                                                        LLAQAGEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEI----KTPTLDKLAAEGVKL 113
                                                                                                                                                                                                   NRWHNERITSSTGKSVWLFNITADPYERVDLSNRYPGIVKKLLRRLSQFNKTAVPVRYPP 537
                                                                                                                           Gaps
                                                                                                                                               2 APRGCAGHPPP----PSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEGGA 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sulfatase) (Steryl-
                                                                                                                                                                                                                                                                                                                                        ALSPDRSEYIKQLPFSKDDVHAVRGGEQQAIADITPKYMEDLDQRWMDYGVKFLDKMAKS
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mus musculus (Mouse).
Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBL_maxID=10090;
                                                                                                                                                                    -------ODKETQQ
                                                                                                                                                                                                                                                                                                                                                            TKPIFLYIAYQAVHSPLQAPGRYFEHYRSI - - - ININRRRYAAMLSCLDEAINNVTLAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DP----IYTKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGNPGYSDWV-PPQSFSNLGP
                                                                                                                                                                                                                                                                                                                                                                                                     KTYGFYNNSIIIYSSDNGCQ----PTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKG
                                                                                                                          97; Mismatches 205; Indels 152;
                                                                                                      Length 551;
                                                             -> I (IN REF. 2).
A04C8BBE5E63E988 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Steryl-sulfatase precursor (EC 3.1.6.2) (Steroid sulfate sulfohydrolase) (Arylsulfatase C) (ASC).
                                                                                                                                                                                                                                                                                                                   AWDYDNGIYSTQMYTQR-----VQQILASHNP----
                                                                                                                                                                  ARKGFAGYDHPNOYLVKPATTIADNMMPVMQHPA-----
                                                                                                     Score 394; DB 1;
Pred. No. 1.7e-23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  624 AA
                                                  POTENTIAL.
       Pfam; PF00884; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
InterPro; IPR000917; Sulfatase.
                                       Hydrolase; Complete proteome.
ACT_SITE 190 190 ECONFLICT 191 191 N
                                                                      60717 MW;
                                                                                                    13.18; 24.58;
                                                                                                                         Matches 147; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                      551 AA;
                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            545 R 545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STS_MOUSE
P50427;
                                                                      SEQUENCE
                                                                                                     Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                       230
                                                                                                                                                                  33
                                                                                                                                                                                       58
                                                                                                                                                                                                                               114
                                                                                                                                                                                                                                                                                                                   229
                                                                                                                                                                                                                                                                                                                                                           556
                                                                                                                                                                                                                                                                                                                                                                                290
                                                                                                                                                                                                                                                                                                                                                                                                     312
                                                                                                                                                                                                                                                                                                                                                                                                                          341
                                                                                                                                                                                                                                                                                                                                                                                                                                              369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STS_MOUSE
SO FFF SO S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                               ò
                                                                                                                                                                  음
                                                                                                                                                                                       ò
                                                                                                                                                                                                         셤
                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                    요
                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                         ď
                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hydrolase; Microsome; Transmembrane; Glycoprotein; Steroid metabolism;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               68 PSTTSTSQPHLIFILADDQGFRDVGYHGSE-IKTPTLDKLAAAEGVKL-ENYYVQPICTPS 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       126 RSQFITGKYQIHTGL-QHS----IIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VAQRPTDYGG------LTRRLADEAALFLRRNRARPFLLFLSFLHVHT------AH 301
                                                                                                                                                                                                                                                                                                                   SEVERAL MEMBRANE-SPANNING DOMAINS THAT COULD SERVE TO ANCHOR THE PROTEIN IN THE MICROSOMAL MEMBRANE.
SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -!- CATALYTIC ACTIVITY: 3-beta-hydroxyandrost-5-en-17-one 3-sulfate H(2)0 = 3-beta-hydroxyandrost-5-en-17-one + sulfate.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LAALGASLAAMAAARWAGLARVPGWALAGTAAAMLAVGGPRSASCLGFRPANCFLMDDLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----DYDNGIYSTQMYTQRVQQ---ILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181 FYRKECM-----PTRRGFDTFFG-----SLLGSGDYYTHYKCDSPGM----
                                                                                                                                                                                                                                                                                           SEQUENCE SHOWS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LINKED (GLCNAC. . ) (POTENTIAL).
025C0CF8659D0A9E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           161;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1; Length 624;
MEDITRE-96241577; pubMed-8673109; Salido E.C., Li X.M., Yen P.H., Martin N., Mohandas T.K., Salido E.C., Li X.M., Yen P.H., Martin N., Mohandas T.K., "Cloning and expression of the mouse pseudoautosomal steroid sulphatase gene (Sts)."; Nat. Genet. 13:83-86(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13.0%; Score 392.5; DB 1; Length 6 24.2%; Pred. No. 2.7e-23; ive 72; Mismatches 184; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STERYL-SULFATASE.
2-AMINO-3-OXOPROPIONIC ACID
                                                                                                                                                                                                                                                                                              THE
                                                                                                                                                                                                                                                                                           SUBCELLULAR LOCATION: MICROSOMAL MEMBRANE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MGD; MGI:99438; Sts.
InterPro; IPR000917; Sulfatase.
Pfam; PF00884; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              66590 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; U37545; AAB09308.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---CGYDLYENDNAAW----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 24.29
Matches 133; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           144
205
234
55
465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22
624
83
                                                                                                                                                                                         DURING PREGNANCY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          55
465
624 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HSSP; P15848; 1FSU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Signal.
```

3iochemistry 37:3654-3664(1998)

```
FADPGFAGRSL----HGAYGDSVEEMDWGVGRVLAALDELGLARETLVYFTSDHGAHVEE 357
                                                                                  LAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPIYTKAKNGSWAAGYGIWNTAIO 446
                                                                                                                                  SAIRVQHWKLLTGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGKS-----VWLFNIT 499
                                                        LGPRGERMGGSNGVFRGGKGNNWEGGVRVPCLVRWPRELSPGRVVAEPTSLMDVFPTVAR 417
                                                                                                                                                      -----NFAPAGANGCFSTHVCLCAGPAHVTAHDPPLLFDLT 513
                                                                                                                                                                                                                                                                                                      01-FEB-1991 (Rel. 17, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
Arylsulfatase A precursor (EC 3.1.6.8) (ASA) (Cerebroside-sulfatase).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     J.C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PARTIAL SEQUENCE, AND 2-AMINO-3-OXOPROPIONIC ACID MODIFICATION SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Lukatela G., Krauss N., Theis K., Selmer T., Gieselmann V., von Figura K., Saenger W.:
"Crystal structure of human arylsulfatase A: the aldehyde function and the metal ion at the active site suggest a novel mechanism for sulfate ester hydrolysis.";
                                  ----PTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLIS
                                                                                                       LAGAELPGDRVIDGRDLMPLLRGDAQRSEHEFLFHYCNAYLQAVR--WHNG-----
                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Schmidt B., Selmer T., Ingendoh A., von Figura K.;
"A novel amino acid modification in sulfatases that is defective multiple sulfatase deficiency.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Fujii T., Kobayashi T., Honke K., Gasa S., Ishikawa M., Shimizu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pohlmann R.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   "Proteolytic processing of human lysosomal arylsulfatase A.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Adams M.D., Kerlavage A.R., Fuldner R.A., Phillips C.A., Ve "Complete sequence of a chromosome 22q subtelomeric BAC."; Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Stein C., Gleselmann V., Kreysing J., Schmidt B., Pc
Wahbed A., Meyer H.E., O'Brien J.S., von Figura K.;
"Cloning and expression of human arylsulfatase A.";
J. Biol. Chem. 264:1252-1259(1989).
                                                                                                                                                                                                                                                                    507 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 19-33 AND 434-479, AND SUBUNITS.
                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-90361046; PubMed-1975241;
Kreysing J., von Figura K., Gleselmann V.;
"Structure of the arylaulfatase A gene.";
Eur. J. Blochem. 191:627-631(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Biochim. Biophys. Acta 1122:93-98(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
MEDLINE=98191325; Pubmed=9521684;
                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
MEDLINE-89093115; PubMed-2562955;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=95354208; PubMed=7628016;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PubMed-1352993
                                                                                                                                                                                                                                                                                          (Rel. 14, Created)
                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Cell 82:271-278(1995).
                                                                                                                                                       SAV----WKAFYFTP
                                                                                                                                                                                                                                                                                                                                                       Homo sapiens (Human)
                                                                                                                                                                                509
                                                                                                                                                                                                      514 RDPGERRPLT 523
                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-92338230;
                                                                                                                                                                               500 ADPYERVDLS
                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                         01-APR-1990
                                                                                                                                                                                                                                                                    ARSA_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Makita A
                                                                                                                                                                                                                                                        ARSA_HUMAN
                                                                                387
                                                                                                        418
                                                                                                                                447
                                                                                                                                                        467
                                                                                                                                                                                                                                                                               g
                                                                                                                                                       g
                                                                                                                                                                                δ
                                                                                                                                                                                                      a
                                  ð
                                                                               õ
                                                                                                                                ò
```

```
"Molecular basis of different forms of metachromatic leukodystrophy.";
New Engl. J. Med. 324:18-22(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                of
                                                                                                                                                                                                   VARIANT MLD GLN-84.
MEDLINE-92344341; Pubmed-1353340;
Kappler J., von Figura K., Gieselmann V.;
"Late-onset metachromatic leukodystrophy: molecular pathology in two
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polten A., Fluharty A.L., Fluharty C.B., Kappler J., von Figura K., Gieselmann V.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARIANT MLD MET-274.

MEDLINE-94004907; PubMed-8104633;

Harvey J.S., Nelson P.V., Carey W.F., Robertson E.F., Morris C.P.;

"An arylsulfatase A (ARSA) missense mutation (T274M) causing late-
infantile metachromatic leukodystrophy.";

Hum. Mutat. 2:261-267(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    caused by substitution
                         MEDLINE-95170731; PubMed-7866401;
Gleselmann V., Zlotogora J., Harris A., Wenger D.A., Morris C.P.;
"Molecular genetics of meterchromatic leukodystrophy.";
Hum. Mutat. 4:233-242(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  causing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Kreysing J., Bohne W., Bosenberg C., Marchesini S., Turpin J.C., Baumann N., von Figura K., Gleselmann V.; High residual arylsulfatase A (ARSA) activity in a patient with late-infantile metachromatic leukodystrophy."; Am. J. Hum. Genet. 53:339-346(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hasegawa Y., Kawame H., Eto Y.; "Mutations in the arylsulfatase A gene of Japanese patients with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARIANT MLD SEK-114.
MEDLINE-94063853; PubMed-7902317;
MEDLINE-94063853; PubMed-7902317;
MONE K., Kobayashi T., Fujii T., Gasa S., Xu M., Takamaru Y.,
Kondo R., Tauji S., Makita A.;
"An adult-type metachromatic leukodystrophy caused by substitu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   "Identification of a mutation in the arylsulfatase A gene of patient with adult-type metachromatic leukodystrophy."; Am. J. Hum. Genet. 48:971-978(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      causes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-91328147; PubMed-1678251;
Gieselmann V., Fluharty A.L., Toennesen T., von Figura K.;
Mutations in the aryisulfatase A pseudodeficiency allele
metachromatic leukodystroppy.";
Am. J. Hum. Genet. 49:407-413(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-91206410; Pubmed-1673291;
Kondo R., Wakamatsu N., Yoshino H., Fukuhara N., Miyatake
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Kafert S., Heinisch U., Zlotogora J., Gleselmann V.; Ma missense mutation P136L in the arylsulfatase A gene or instability and loss of activity of the mutant enzyme."; Hum. Genet. 95:201-204(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARIANT MLD LEU-426, AND VARIANTS CYS-193 AND SER-391.
MEDLINE-91074201; Pubmed-1670590;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 serine for glycine-122 in arylsulfatase A. Hum. Genet. 92:451-456(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VARIANT MLD ARG-245.
MEDLINE-93319632; Pubmed-8101083;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VARÍANT MLD SER-309,
MEDLINE-93318834; PubMed-8101038;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=95163939; PubMed=7860068;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           metachromatic leukodýstrophy.";
DNA Cell Biol, 12:493-498(1993).
                                                                                                                                                                                                                                                                                                                                                                                Ann. Neurol. 31:256-261(1992).
REVIEW ON MLD VARIANTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARIANT MLD SER-122.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARIANT MLD LEU-136.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         VARIANT MLD PHE-96.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VARIANT MLD ASP-99.
                                                                                                                                                                                                                                                                                                                                                   siblings."
```

Wed Jul 24 10:05:46 2002

```
Ricketts M.H., Poretz R.D., Manowitz P.;
"The R496H mutation of arylsulfatase A does not cause metachromatic
 MEDLINE-98415722; PubMed=9744473;
   RX
RA
                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                        à
                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ļ'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VARIANT MLD 406-SER--THR-408 DEL.
MEDLINE-98141126; PubMed-9490297;
Regis S., Filocamo M., Stroppiano M., Corsolini F., Caroli F.,
Gatti R.;
A 9-bp deletion (2320del9) on the background of the arylsulfatase A
pseudodeficiency allele in a metachromatic leukodystrophy patient and
in a patient with nonprogressive neurological symptoms.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Characterization of two arylsulfatase A missense mutations D335V and T274M causing late infantile metachromatic leukodystrophy."; Hum. Mutat. 7:311-317(1996).
                                              Gieselmann V., Polten A., Kreysing J., von Figura K.;
"Arylsulfatase A pseudodeficiency: loss of a polyadenylylation signal
                                                                                                                                                                                                                                                                                                                                                                           VARIANTS MLD LEU-82; TYR-172; CYS-201; GLN-311; VAL-335 AND TRP-390. MEDILNE-96047150; PubMed-7581401; Barth M.L., Fensom A., Harris A.; "Identification of seven novel mutations associated with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Draghia R., Letourneur F., Drugan C., Manicom J., Blanchot C., Kahn A., Poenaru L., Caillaud C.; "Metachromatic leukodystrophy: identification of the first deletion exon 1 and of nine novel point mutations in the arylsulfatase A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE=98260871; PubMed=9600244;
Gomez-Lira M., Perusi C., Mottes M., Pignatti P.F., Manfredi M.,
Rizzuto N., Salviati A.;
Molecular genetic characterization of two metachromatic
leukodystrophy patients who carry the T799G mutation and show
different phenotypes; description of a novel null-type mutation.";
Hum. Genet. 102:459-463(1998).
                                                                                                                                                                                                                                                                                       Barth M.L., Fensom A., Harris A.;
"Missense mutations in the arylsulphatase A genes of metachromatic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-98112481; PubMed-9452102;
Coulter-Mackle M.B., Gagnier L.;
"Two novel mutations in the arylsulfatase A gene associated with
juvenile (R390Q) and adult onset (H397Y) metachromatic
                                                                                                                           VARÍANT MLD LEU-426.
MEDLINE-93202658; PubMed-8095918;
Barth M.L., Fensom A., Harris A.;
"Prevalence of common mutations in the arylsulphatase A gene in metachromatic leukodystrophy patients diagnosed in Britain.";
Hum. Genet. 91:73-77(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHARACTERIZATION OF VARIANTS MET-274 AND VAL-335.
MEDLINE-96303701; PubMed-8723680;
Hess B., Kafert S., Heinisch U., Wenger D.A., Zlotogora J.,
                                                                                and N-glycosylation site.";
Proc. Natl. Acad. Sci. U.S.A. 86:9436-9440(1989)
                                                                                                                                                                                                                                                      VARIANTS MLD VAL-212; VAL-224 AND TYR-295.
MEDLINE=94154687; PubMed=7906588;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mutat. Suppl. 1:S254-S256(1998)
                                                                                                                                                                                                                                                                                                                         leukodystrophy patients.";
flum. Mol. Genet. 2:2117-2121(1993)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=97245886; PubMed=9090526;
                              MEDLINE-90083282; PubMed=2574462;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARIANTS MLD PRO-135 AND SER-179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARIANTS MLD GLN-390 AND TYR-397
                                                                                                                                                                                                                                                                                                                                                                                                                                               metachromatic leukodystrophy.";
Hum. Mutat. 6:170-176(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mutat. 9:234-242(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Genet. 102:50-53(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARIANT HIS-496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gieselmann V.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VARIANTS MLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HEM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hum.
```

```
19;
                                                                                                                                                                                                            278
                                                                                                                                                                                                                                                                                                                                                                                         335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FTQGSAHSD-----TTADPACHASSSLTAH--EPPLLYDLSKDPGENYNLLGGVAGA 448
                                                                                                                                                                                                                                                                        FDTFFGSLLGSGDYYTHYK------CDSPGMCGYDLYENDNAAWDYDNGIYS 238
                                                                                                                                                                                                                                                                                                                  -----DQGLVP 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        336 GSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDED 395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    293 GCSGLLRCGKGTTYEGGVREPALAFWPGHIAPG-VTHELASSLDLLPTLAALAGAPL-PN 350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               396 IQLDGYDIWETISEGLRSPRVDILHNIDPIYTKAKNGSWAAGYGIWNTAIQSAIRVQHWK 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     456 LLTGNPCYSDWVPPQSFSNLGPNRWHNERITSSTGKSVWLFNITADPYERVDLSNRYPGI 515
                                                                                              76 PHLIFILADDQGFRDVGYHG-SEIKTPTLDKLAAEGVKLENYYVQ-PICTPSRSQFITGK 133
                                                                                                                                                                                 134 YQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKEC-MPTRRG 192
                                                    Gaps
                                                                                                                       239 TQM------YTQRVQQILA-SHNPTKPIFLYIAYQAVHSPLQAPGRY
                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEHYRSIININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNGGQP---TAG
                                                  92;
      Length 507;
12.5%; Score J.C., 2.8e-22; 26.1%; Pred. No. 2.8e-22; +ive 75; Mismatches 195; Indels
                                                                                                                                                                                                                                                                                                                    141 FHRFLGI-----PYSHDQGPCQNLTCFPPATPCD--GGC----
                             Best Local Similarity 26.13
Matches 128; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            516 VKKLLRRLSQ 525
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   449 TPEVLQALKQ 458
           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                   279
                                                                                                                                                                                                                                                                        193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              398
```

Search completed: July 23, 2002, 19:07:50 Job time: 366 sec

•

•

Wed

```
Drosophila melanog
Human biallelic po
Drosophila melanog
Drosophila melanog
Drosophila melanog
Gene encoding a su
Human pRO708 hybri
Human pRO708 hybri
Human pRO708 hybri
Human jmmune syste
Drosophila melanog
Drosophila melanog
                                                                                                                                                                                                                                       Hepatitis C virus
Human secreted pro
Human polynucleoti
                                           Human colon cancer
Drosophila melanog
Human colon cancer
                                                                                                                                                                                                                                                                                   Human immune syste
Human metastasis a
                                                                                                                                                                                                                                                                                                                                                            Human secreted pro
Human secreted pro
Human PRO6006 cDNA
                                                                                                                                                                                                                               Human immune syste
                                                                                                                                                                                                                                                                         Human immune syste
                                                                                                                                                                                                                                                                                                       Human polynucleoti
Human immune syste
Human 25278 sulfat
                                                                                                                                                                                                                                                                                                                                        Human inflammatory
                                                                                                                                                                                                                                                                                                                                                   Human inflammatory
                                                                                                                                                                                                                                                                                                                              cDNA encoding for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sulfatase; 23553 sulfatase; nootropic; neuroprotective; antibacterial; antlinflammatory; vasotropic; antitumor; gene therapy; human; ss.
                                                                                                                                                                                                                                                                                                                                                              Human
Human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Tsai F;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human 23553 sulfatase polypeptide encoding cDNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Rudolph-Owen LA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /*tag= a
/product= "23553 sulfatase"
                                                                                                                                                                                                                                                                                                                                                                                                                  ALIGNMENTS
                              AAS28640
AAA02370
AAA00278
AAA00278
AAX11642
AAX12813
ABL09675
AAX12813
ABL12894
AAX13893
AAX333990
AAC78675
                                                                                                                                                                                 ABL33354
ABL12885
ABL12858
ABL09674
                                                                                                                                                                                                                            ABL33578
AAT89608
AAC59201
AAI87912
ABL32171
ABL33082
                                                                                                                                                                                                                                                                                           ABL34528
AA180370
ABL32681
AA8229116
AAH92784
AAV92785
AAC91487
AAC5528
           AAS41205
AAH99356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers
324..2123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAH46863 standard; cDNA; 2253 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Glucksmann MA, Williamson M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MILL-) MILLENNIUM PHARM INC
  31-JAN-2001; 2001WO-US03266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    31-JAN-2000; 2000US-0495823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25-SEP-2001 (first entry)
uuu
onwaaanu444auuuuuuuuuuuuuuuuuuuuuuu
u4amuu/40muon4mumumuuu11110000000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO200155411-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           02-AUG-2001
AAH46863;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAH46863
 RESULT
                                                                                                                     O O
                                                                                                                                                                                                                       0 0
                                                                                                                                                                                                                                                                         U
                                                                                                                                                                                                                                                                                                                    O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 July 23, 2002, 15:05:29 ; Search time 356.89 Seconds
(without alignments)
10901.197 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DNA encoding
PRO708 (UNQ3
PRO708 nucle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23553 sulfat
CDNA encodin
CDNA encodin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                cancer
sulfat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Human 23553 sulfat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                       **SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*

$\times_{\text{SIDS1}} \text{Gradata/geneseq} \text{Geneseqn-embl/NA1981.DAT:*}

$\times_{\text{SIDS1}} \text{Gradata/geneseqqn-embl/NA1981.DAT:*}

$\times_{\text{SIDS1}} \text{Gradata/geneseqq/geneseqn-embl/NA1981.DAT:*}

$\times_{\text{SIDS1}} \text{Gradata/geneseqq/geneseqn-embl/NA1981.DAT:*}

$\times_{\text{SIDS1}} \text{Gradata/geneseqq/geneseqn-embl/NA1991.DAT:*}

$\times_{\text{SIDS1}} \text{Gradata/geneseqq-embl/NA1991.DAT:*}

$\times_{\text{SIDS1}} \text{Gradata/geneseqq-embl/NA2001.DAT:*}

$\times_{\text{SIDS1}} \text{Gradata/geneseqq-embl/NA2001.DAT:*}

$\times_{\text{SIDS1}} \text{Gradata/geneseqq-embl/NA2001.DAT:*}

$\times_{\text{SIDS1}} \text{Gradata/geneseqq-embl/NA20
                                                                                                                                       1 cacgegteegeecaegegte......tgeeaeetggtgeegaatte 2266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                colon
25278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Human
                     Compugen Ltd
                                                                                                                                                                                                              Total number of hits satisfying chosen parameters:
                                                                                                                                                                                            1736436 seqs, 858457221 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUMMARIES
          GenCore version
Copyright (c) 1993 - 2000
                                                                                                                                                                                                                                                                    Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAH46863
AAC78483
AAC33987
AAC33987
AAC32726
AAS22490
AAAO2344
AAAO2344
                                                    nucleic search, using sw model
                                                                                                                                                            IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
                                                                                                                                                                                                                                       seq length: 0
seq length: 2000000000
                                                                                                             US-09-495-823-8
2266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2253
4639
4640
4650
11800
1375
1340
729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98.6
89.8
89.8
89.8
89.8
32.3
27.6
26.3
```

Score

Result Š 2234.2 2033.8 2033.8 2033.8 1781.6 890 736.2 625.2

Post-processing:

Database

Minimum DB Maximum DB

Perfect score:

Sequence:

٠

OM nucleic

ö

Scoring table:

Searched:

P-PSDB; AAB85483 

2001-476214/51

acute Novel human sulfatase polypeptides useful for treating and diagr sulfatase-related disorders such as cerebrovascular diseases, ac meningitis, multiple sclerosis, degenerative diseases and tumor Novel human sulfatase polypeptides sulfatase-related disorders such as

English Fig 15; 180pp; Claim 2;

The invention provides 22438, 23533, 25278 or 26212 human sulfatase polypeptides and polynucleotides. The sulfatase genes and polypeptides are useful for treating disorders involving the brain such as cerebrovascular diseases, infections such as acute meningitis, demyelinating diseases including multiple sclerosis, degenerative diseases affecting the cerebral cortex including Alzheimer's disease and Pick disease, spinocerebellar degenerations including spinocerebellar degenerative diseases affecting motor neurons including amyotrophic lateral sclerosis, inborn errors of metabolism such as leukodystrophies, toxic and acquired metabolic diseases, including vitamin deficiencies, and neurocutaneous syndromes (phakamatoses) including neurofibromotosis. sulfatase polypeptide human 23553 ø represents present sequence encoding

T; 0 other; 540 .; G 554 ပ 2253 BP; 624 A; 535 Seguence

ij 300 360 ggaagctgctctagggaaggggggaggaggaggagaaaagtgaaatgtgctggagaagagcg 120 360 480 480 Gaps è0 9 agteteacetgggacetgagtgaatggeteceaggggetgtgeggggeteegeeteege agccctccttgttcttccggagtcccatccattaagccatcacttctggaagattaaagt accacggatctgagattaaaacacctactcttgacaagctcgctgccgaaggagttaaac tgtcggacatggtgacagctgagaggagaggagtttcttgccaggtggagagtcttca agtotoacctgggacctgagtgaatggctcccagggggctgtgcgggggcatccgcctccgc aggaaggaggccttactagctcaagctggagagaaactagagcccagcacaacttcca cctcccagccccatctcatcctagcggatgatcagggatttagagatgtgggtt tctggatcctctgcctcacttatggttacctgtcctggggccaggccttagaagagg 1; Length 2253; 4; Indels DB 22; Score 2234.2; Pred. No. 0; 1; Mismatches 98.68; 99.78; Query Match 98.6 Best Local Similarity 99.7 Matches 2247; Conservative 181 Н Н 121 181 241 61 61 121 241 301 301 361 421 121 481 361 481 541 601 ò 셤 ò g ò g ò q ŏ qq ò g à q g g Q ò à ò ò

1020 1500 099 720 780 780 840 840 900 900 960 tacctctggacaatgccaccctacctcagaaactgaaggaggttggatattcaacgcata atgetgecatgettecetgettagatgaageaateaacaaegtgacattggetetaaaga cggcaggaaggagtaactggcctctcagaggtagcaaaggaacatattgggaaggaggga accacggatctgagattaaaaacacctactcttgacaagctcgctgccgaaggagttaaac agtatcagatacacaccggacttcaacattctatcataagacctacccaacccaactgtt ttgataccttttttggttcccttttgggaagtggggattactatacacactacaaatgtg ttgataccttttttggttcccttttgggaagtggggattactatacaccactacaatgtg atgaggacattcaactagatggctatgatatctgggagaccataagtgagggtcttcgct cctgggcagcaggctatgggatctggaacactgcaatccagtcagccatcagagtgcagc tggagaactactatgtccagcctatttgcacaccatccaggagtcagtttattactggaa atgctgccatgcttcctgcttagatgaagcaatcaacaacgtgacattggctctaaaga cttatggtttctataacaacagcattatcatttactcttcagataatggtggccagccta tccgggctgtaggctttgtgcatagcccacttctgaaaaacaagggaacagtgtgtaagg aacctgtgcacatcactgactggtaccccactctcatttcactggctgaaggacagattg cctgggcagcagctatgggatctggaacactgcaatccagtcagccatcagagtgcagc actggaaattgcttacaggaaatcctggctacagcgactgggtcccccctcagtcttca 841 1081 1141 1141 1381 1381 1021 1081 1201 1201 1261 1261 1321 1441 1501 1501 1561 1621 1681 661 199 721 721 781 781 841 901 901 961 961 1021 1321 1441 1561 1681 δy qq qq В qq q q ò q ò q q g Q g g g QQ QQ q g δ ò ò οy ò ò ò ŏ ò à à ò ò δý ò

```
PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep; ss; dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha; blood; chondrocyte cell; cell proliferation; cell differentiation; colon; adrenal; lung; breast; prostate; rectum; cervix; liver; qenetic disorder:
                                                                                                                                                                                                                               gcaacctgggaccgaaccggtggcacaatgaacggatcacctcgtcaactggcaaaagtg 1800
                 1741 gcaacctgggaccgaaccggtggcacaatgaacggatcaccttgtcaactggcaaaagtg 1800
                                                                                                                tcaggtatccccccaaagaccccagaagtaaccctaggctcaatggaggggtctagggac 1980
                                                                                                                                                                                           aaaagaaaagc.aaaaaagaagaagaagaaacagcagaaagcagtctcaaggttcaacttgcc 2099
                                                                                                                                                                                                             aaaagaaaagcaaaaaaaaaagaagaaacagcagaaagcagtctcaggttcaacttgcc
                                                                                                                                                      catggtatagagagaaaccaagaaaaagaaagccaagcaaaaatcaggctgagaaaagc
                                                                                                                                                                                                                                                                                                                                                                                                                                   Human DNA encoding PRO polypeptide sequence #38.
                                                                                                                                                                                                                                                                                                         AAS45962 standard; cDNA; 4639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2000WO-USO5841
2000US-18-02P
2000US-18-93-2P
2000US-18-93-2B
2000WO-USO6884
2000US-19-08-2P
2000US-19-07-P
2000US-19-07-P
2000US-19-19-07-P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28-FEB-2001; 2001WO-US06520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2000WO-US05601
                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200168848-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  02-MAR-2000;
03-MAR-2000;
06-MAR-2000;
14-MAR-2000;
114-MAR-2000;
21-MAR-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21-MAR-2000;
21-MAR-2000;
21-MAR-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAR-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                 18-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              primer
                                                                                                                                                                                                                                                                                                                                                                                               AAS45962;
                                                                                                                                                                                                                                                 2101
                                                                                                                                                                                                                                                                                       2161
                                                                                                                                                                                                                                                                                                                            2221
1741
                                                      1801
                                                                         1861
                                                                                            1861
                                                                                                               1921
                                                                                                                                                    1981
                                                                                                                                                                                          2041
                                                                                                                                                                                                            2041
                                                                                                                                                                                                                               2100
                                                                                                                                                                                                                                                                    2160
                                    1801
                                                                                                                                 1921
                                                                                                                                                                      1981
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PCR
                                                                                                                                                                                                                                                                                                                                                                 AAS45962
                g
                                                     a
                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                        q
                                                                         ò
                                                                                           g
                                                                                                                               g
                                                                                                                                                    ŏ
                                                                                                                                                                     g
                                                                                                                                                                                                           g
                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                        à
```

```
primers for PRO POLYPEPTIGES of the invention. The sequences of the invention can be used to detect the presence of a tumour in a mammal by comparing the level of expression of a PRO polypeptide in a test sample of cells from the animal and a control sample of normal cells, whereby a higher level of expression in the test sample indicates the presence of a tumour in the mammal is include dogs, cars, cattle, horses, sheep, pigs, goats and rabbits but are preferably human. The polypeptides can be used to stimulate tumour necrosis factor (TNF) alpha release from human to be used to stimulate the proliferation or differentiation of chondrocyte cells. The PRO proteins can be used to determine the presence of tumours and also susceptibility to tumour development, particularly adrenal, lung, colon, creast, prostate, rectal, cervical, or liver tumours, in mammalian subjects. The oligonucleotide probes specific for the PRO nucleic acids can be used for genetic analysis of individuals with genetic disorders.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gurney AL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Novel nucleic acids encoding PRO polypeptides, used to diagnose the presence of tumours, such as prostate and breast tumours, in mammals to screen for modulators of the compounds -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 22; Length 4639;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Godowski PJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 4639 BP; 1426 A; 956 C; 1025 G; 1232 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              , Goddard A, Go
Wood WI, Zhang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 2033.8;
Pred. No. 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŗ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 2; Fig 75; 774pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Watanabe CK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Desnoyers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   89.8%;
99.0%;
                         2000US-1930332P.
2000US-193053P.
2000WO-US08439.
2000US-194449P.
2000US-194647P.
2000US-195975P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  2000WO-US14042.
2000WO-US14941.
2000WO-US15264.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2000US-0644848.
2000WO-US23328.
2000WO-US30952.
                                                                                                                                                                          2000US-196000P.
2000US-196187P.
                                                                                                                                                                                                                                                                                                 2000US-198585P.
2000US-199397P.
                                                                                                                                                                                                                                                                                                                                                  2000US-199550P.
2000US-199654P.
                                                                                                                                                                                                                         2000US-196690P
2000US-196820P
                                                                                                                                                                                                                                                                            2000US-198121P
                                                                                                                                                                                                                                                                                                                                                                                                   2000US-201516P
2000WO-US13705
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2000US-209832P
2000WO-US20710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 99.0
Matches 2066; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GETH ) GENENTECH INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2001-602746/68.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Chen J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Smith V,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             P-PSDB; AAU29061
28-MAR-2000; 29-MAR-2000; 29-MAR-2000; 20-MAR-2000; 20-MA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Baker KP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pan J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Dρ
     ò
```

ä

82

ŏ

4

8 8 7 7 9 9 0 0	48 00 01 61	1128 621 1188 681 1248	1308 801 1368 861 1428 921	981 1548 1041 1608 1101 1668 1151 1728 1221
ggaggaggaggagaatgtgctggagaagggagcctccttgttcttccgga gtcccatccattaagccatcacttctggaagattaaagttgtcggacatggtgacagctg lilililililililililililililililililili	gaatggctcccaggggctgtgcggggcatccgcctccgccttctccacaggcctgtgtc gtcctggaaagatgctagcaatgggggcgctggcaggattctggatcctctgcctcctc	ctcaagctggagagaaactagagcccagcacaacttccacaccccactctcatt.  tcatcctagcggatgatcagggatttagagatgtgggttaccaggatctgagattaaa.  [	ctattigcacaccatccaggagtcagtttattactggaaagtatcagatacaccggga ttcaacattctatcatcagaacctacccaaccca	ttttgggaagtggggttactatacacactacaaatgtgacagtcctgggatgtgtgct liiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii
4 4 0 0 0 0 0 0	0 0 4 4 0 0 1	1069 562 1129 622 1189	1249 742 1309 802 1369 862 862	922 1489 982 1549 1042 1609 1102 1162
			2	

1848 1908 1401 1968 1461 2028 2088 1581 2148 1641 2207 1701 2267 1761 2327 1821 2387 1881 2447 1941 2507 2001 2567 2060 2627 1521 Human; secreted protein; transmembrane protein; PRO; EST; cytostatic; tagatgaagcaatcaacaacgtgacattggctctaaagacttatggttctataacaaca ggtaccccactctcátttcactggctgaaggacagattgatgaggacattcaactagatg ggtaccccactctcatttcactggctgaaggacagattgatgaggacattcaactagatg ggcacaatgaacggatcacctcgtcaactggcaaaagtgtatggcttttcaacatcacag ggcacaatgaacggatcaccttgtcaactggcaaaagtgtatggcttttcaacatcacag tacggaggeteteacagtteaacaaaactgeagtgeeggteaggtateceecaaagaee ccagaagtaaccctaggctcaatggaggggtctagggaccatggtatagagaggaaacca (UNQ372) nucleotide sequence SEQ ID NO:113 ВР 4640 (first entry) CDNA; standard; Human PRO708 08-FEB-2001 AAC78483 AAC78483; 3 1942 1642 2002 2061 1222 1789 1282 1849 1342 1909 1402 1969 1462 2029 1522 2089 1582 2149 2208 1702 2268 1762 2328 1822 2388 1882 2448 2508 2568 2628 AAC78483 RESULT Dp Q g g ద g Q οy q g ò οp g ò qq οy pp δ 임 ò qq δ οy δ ò ò ò ò ò

us-09-495-823-8.rng

2

Page

201

```
gtcccatccattaagccatcacttctggaagattaaagttgtcggacatggtgacagctg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1042
                                                                                                                                                                                                                    1009
                                                                                                                                                                                                                                                           1069
                                                                                                                                                                                                                                                                                                1129
                                                                                                                                                                                                                                                                                                                                       1189
                                                                                                                                                                                                                                                                                                                                                                            1249
                                                                                                                                                                                                                                                                                                                                                                                                                  1309
                                                                                                                                                                                                                                                                                                                                                                                                                                                         1369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1609
 142
                   709
                                       202
                                                         691
                                                                            262
                                                                                               829
                                                                                                                   322
                                                                                                                                      889
                                                                                                                                                          382
                                                                                                                                                                              949
                                                                                                                                                                                                 44.2
                                                                                                                                                                                                                                       502
                                                                                                                                                                                                                                                                              562
                                                                                                                                                                                                                                                                                                                     622
                                                                                                                                                                                                                                                                                                                                                          682
                                                                                                                                                                                                                                                                                                                                                                                                 742
                                                                                                                                                                                                                                                                                                                                                                                                                                       802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          982
                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QQ
                                                                                               g
                                                                                                                                     g
                                                                                                                                                          δy
                                                                                                                                                                             셤
                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                       δλ
                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                               Dp
                                                                                                                                                                                                                                                                                                                   οχ
                                                                                                                                                                                                                                                                                                                                     Ω
                                                                                                                                                                                                                                                                                                                                                         δy
                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
               g
                                       οy
                                                        9
                                                                            οy
                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAC78458 to AAC78599 represent polynucleotide and EST (expressed sequence tag) sequences which encode secreted or transmembrane PRO polypeptides. The PRO polynucleotides and polypeptides have cytostatic activity. The polynucleotides and polypeptides can be used for detecting the presence of PRO polypeptides in samples, for linking bioactive molecules to cells and for modulating biological activities of cells, using the polypeptides for specific targeting. The polypeptide targeting can be used to kill the target cells, e.g. for the treatment of cancers. The polypeptide pairs provide specific targeting of bioactive molecules to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
                                                                                                                                                                                                                                                                                                                                                                                                s used in detection methods, cells, and to modulate
                                                                                                                                                                                                                                                                                                                     Ā
                                                                                                                                                                                                                                                                                                          Eaton DL;

1, Gerritsen Millan KJ;
ROY MA;
Wood WI;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 4640;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            955 C; 1026 G; 1232 T; 0 other;
                                                                                                                                                                                                                                                                                                          KP, Botstein D, Desnoyers L, Ea
E, Fong S, Gao W, Gerber H, G
Pap, Grimaldi CJ, Gurney AL, Hil
Napier MA, Pan J, Paoni NF, Ro
TA, Tumas D, Williams PM, Wood
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         isolation of the PRO polynucleotide sequences
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB
 SS
                                                                                                                                                                                                                                                                                                                                                                                              Novel PRO polypeptides and polynucleotides to target bloactive molecules to specific coellular activities -
detection; cancer;
                                                                                                                                                                                                                                                                                                                                                                                                                                      Fig 42; 636pp; English.
                                                                                                                                                                                                                           99WO-US31243.
99WO-US31274.
2000WO-US00219.
2000WO-US00277.
2000WO-US00376.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         89.8%;
                                                                                                         99US-0123957.
99US-013032.
99US-0131445.
99US-0134287.
99US-0141037.
99US-0145698.
99US-0145508.
                                                                                                                                                                                                        99WO-US28565,
                                                                            2000WO-US04341
                                                                                                                                                                                                99WO-US28551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BP; 1427 A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 89.8
Best Local Similarity 99.0
Matches 2066; Conservative
                                                                                                                                                                                                                                                                                                          , Baker KP,
Filvaroff E,
                                                                                                                                                                                                                                                                                                                                               Stewart TA,
                                                                                                                                                                                                                                                                                                                           Godowski PJ,
Kuo SS, Na
 tag;
                                                                                                                                                                                                                                                                                       (GETH ) GENENTECH INC.
                                                                                                                                                                                                                                                                                                                                                                  WPI; 2000-611443/58.
P-PSDB; AAB44257.
 expressed sequence
                                     WO200053756-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 4640
                                                                                                                                                                                                                                               05-JAN-2000;
06-JAN-2000;
06-JAN-2000;
                                                                            18-FEB-2000;
                                                                                              08-MAR-1999;
12-MAR-1999;
29-MAR-1999;
21-APR-1999;
                                                                                                                                                                            -OCT-1999;
-NOV-1999;
-DEC-1999;
                                                                                                                                                                                                        -DEC-1999;
                                                                                                                                                                                                                                                                                                          Ashkenazi AJ
                                                                                                                                                                                                                                                                                                                  Ferrara N,
Goddard A,
Kljavin IJ,
                                                                                                                                                                                                                            30-DEC-1999;
30-DEC-1999;
                                                                                                                                                                                                                                                                                                                                                Shelton DL,
                                                                                                                                      28-APR-1999;
                                                                                                                                               14-MAY-1999
                                                                                                                                                          .JUN-1999
                                                                                                                                                                  JUL-1999;
                                                         14-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 2;
                    Homo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         the
                                                                                                                                                                              29-
30-
02-
```

```
1008
                                                                                                                                                                                                                                                                            1068
                                                                                                                                                                                                                                                                                                                                                                                   1188
                                                                                                                                                                                                                                                                                                                                                                                                                                      1248
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1041
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                tagatgaagcaatcaacaacgtgacattggctctaaagacttatggtttctataacaaca 1281
                                                                                                                                                                     948
                                                                                                                                                                                                441
                                                                                                                                                                                                                                                     501
                                                                                                                                                                                                                                                                                                        561
                                                                                                                                                                                                                                                                                                                                                            621
                                                                                                                                                                                                                                                                                                                                                                                                             681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     801
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         861
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           921
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               981
                                                                                                                  888
                                     261
                                                                                        321
                                                                                                                                            381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ttttgggaagtggggattactatacacactacaaatgtgacagtcctgggatgtgtggct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                agatgtacactcagagagtacagcaaatcttagcttcccataaccccacaaagcctatat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ctcaagctggagagaaactagagcccagcacaacttccacctcccagccccatctcattt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        tacctcagaaactgaaggaggttggatattcaacgcatatggtcggaaaatggcacttgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ttttgggaagtggggattactatacacactacaaatgtgacagtcctgggatgtgtggct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   atgacttgtatgaaaacgacaatgctgcctgggactatgacaatggcatatactccacac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                agatgtacactcagagagfacagcaaatcttagcttcccataaccccacaaagcctatat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               aacactaccgatccattatcaacataaacaggaggagatatgctgccatgctttcctgct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        aacactaccgatccattatcaacataaacaggagaagatatgctgccatgcttcctgct
gegecegeaseggegegggggggtggtteteegegtggagteteaeetgggaaettgagt
                                                                                                    cttatggttacctgtcctggggccaggccttagaagaggaggaagaaggaggccttactag
                                                                                                                                                                                                                                                                               cttatggttacctgtcctggggccaggccttagaagaggaggaagaaggggccttactag
                                                                                                                                                                                                                                                                                                        ctcaagctggagagaactagagccagcacaacttccacctcccagcccattt
                                                                                                                                                                                                                                                                                                                                                                                                               cacctactcttgacaagctcgctgccgaaggagttaaaactggagaactactatgtccagc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ctatttgcacaccatccaggagtcagtttattactggaaagtatcagatacacccggac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ttcaacattctatcataagacctacccaacccaactgtttacctctggacaatgccaccc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  tacctcagaaactgaaggaggttggatattcaacgcatatggtcggaaaatggcacttgg
                                       agaggaggaggatttcttgccaggtggagagtcttcaccgtctgttgggttgcatgtgt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gtggagatattaacttttttttttttttttctttggtggaagctgctctaggagggg 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 2033.8;
Fred. No. 0;
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 549
```

qq

ô

g

å

XX OS Homo sapiens.	. PN WO9946281-A2.	PD 16-SEP-1999.	08-MAR-1999; 10-MAR-1998;	PR 11-MAR-1998; 98US-0077632. PR 11-MAR-1998; 98US-0077641. PR 12-MAR-1998; 98US-0077649. PR 12-MAR-1998; 98US-0077791.	13-MAR-1998; 17-MAR-1998; 20-MAR-1998;	20-MAR-1998; 20-MAR-1998;	20-MAR-1998; 25-MAR-1998; 26-MAR-1998;	27-MAR-1998; 27-MAR-1998; 27-MAR-1998;	27-MAR-1998; 27-MAR-1998; 30-MAR-1998;	30-MAR-1998; 31-MAR-1998;	31-MAR-1998; 31-MAR-1998;	31-MAR-1998; 01-APR-1998;	01-APR-1998; 01-APR-1998; 01-APR-1998;	08-APR-1998; 08-APR-1998;	08-APR-1998; 09-APR-1998;	09-APR-1998; 09-APR-1998;	15-APR-1998; 15-APR-1998; 15-APP-1008:	15-APR-1998; 21-APR-1998;	21-APR-1998; 22-APR-1998;	22-APR-1998; 22-APR-1998;	23-APR-1998; 23-APR-1998;	27-APR-1998; 28-APR-1998; 29-APP-1998;	29-APR-1998; 29-APR-1998;	29-APR-1998; 29-APR-1998; 29-APP-1999;	29-APR-1998;	29-APR-1998; 30-APR-1998;	05-MAY-1998;	06-MAY-1998;	07-MAY-1998; 07-MAY-1998; 07-MAY-1998;
	1282 gcattatcatttactcttcagataatggtggccagcctacggcaggaggaggagtaactggc 1341	tcagataatggtggccagcctacggcaggagggagtaactgg	1342 ctctcagaggtagcaaaggaacatattgggaaggagggatccgggctgtaggctttgtgc 1401 	uggaacagtgtgtaaggaacctgtgcacatca 	2 ggtaccccactctcatttcactggctgaaggacagattgatgatgaggacattcaactaga	ggtaccccactctcatttcactggctgaaggacagattgatgaggacattcaactagatg 2	1347 gctargatarctgggagaccataagrgagggtcttcgctcacccgggtagatatttgc 1381 	ataacattgaccccatatacaccaaggcaaaaatggctcctgggcagcaggctatggga	s ataacatugacece-tatacacaaggeaaaaaaaggeteteegggeageageageage 5 tetadaacactacaatecaateaaceateaaaatacaacaetaaaattacttac	tctggaacactgcaatccagtcagccatcagagtgcagcactggaaattgcttacaggaa	ctgggtccccctcagtctttcagcaacctgggaccg		ggcacaatgaacggatcacctcgtcaactggcaaaagtgtatggcttttcaacatcacag 1	ggcacaatgaacggatcaccttgtcaactggcaaaagtgtatggcttttcaacatcacag 23	ccgacccatatgagagggtggacctatctaacaggtatccaggaatcgtgaagaagctcc 188	ccgacccatatgagagggtggacctatctaacaggtatccaggaatcgtgaagaagctcc 2	1882 tacggaggeteteacagtteaacaaactgcagtgccggteaggtatecececaaagace 1941 1111111111111111111111111111111111	2 ccadaaqtaaccctadcctcaatqqaqqqqtctaqqqaccacqqqacqa		aagcaaaaatcaggctgagaaaaagcaaaagaaaagc-aaa		2061 agaagaaacagcagaaagcagtctcaggttcaacttgccattcaggt 2107	aacagcagaaagca	ULF 4 33987	AAZ33987 standard; cDNA; 4650 BP.	AAZ33987;	07-DEC-1999 (first entry)	Human PR0708 nucleotide sequence.	KW Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation; KW probe; blood coagulation disorder; cancer; cellular adhesion disorder; KW secreted protein; transmembrane protein; ss.

gegecegeaseggegeggggggegtggtteteegegtggagteteaeetgggaeetgagt

~

Jul

Wed

```
The present invention describes secreted and transmembrane polypeptides and their polynucleotides. The nucleotide sequences are useful as sources of probes, primers, for chromosome mapping, and for generation of antisense sequences. They can also be used to create transgenic animals. The proteins can be used to treat a variety of diseases and disorders, depending on their function. Diseases that may be treated include blood coagulation disorders, cancers and cellular adhesion disorders. They may also be used to raise antibodies. AAX33381 to AAX34338, and AAX41685 to AAX41774 represent polynucleotide and polypeptide sequence given in the exemplification of the present
                                                                                                                                                                                                                                                                                                                                                                             New secreted and transmembrane polypeptides and their polynucleotides, useful for treating blood coagulation disorders, cancers and cellular adhesion disorders -
                                                                                                                                                                                                                                                                                                                          Ë
                                                                                                                                                                                                                                                                                                                          Chen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   G; 1235 T; 0 other;
                                                                                                                                                                                                                                                                                                                          ΚP,
                                                                                                                                                                                                                                                                                                                           Baker
                                                                                                                                                                                                                                                                                                                          J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C; 1026
                                                                                                                                                                                                                                                                                                                          Yuan
                                                                                                                                                                                                                                                                                                                          À,
                                                                                                                                                                                                                                                                                                                                                                                                                                2; Fig 42; 530pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 4650 BP; 1431 A; 958
98US-0084639.
98US-0084640.
98US-0085333.
98US-0085338.
98US-0085338.
98US-0085579.
98US-0085580.
98US-0085580.
98US-0085580.
98US-0085580.
                                                                                                                                                                980S-0086023.
980S-0086414.
980S-0086414.
980S-0086486.
980S-0087098.
980S-0087106.
980S-0087208
                                                                                                                                                                                                                                                                                                                          Gurney
                                                                                                                                                                                                                                                                           98US-0100038
                                                                                                                                                                                                                                                                                                 (GETH ) GENENTECH INC
                                                                                                                                                                                                                                                                                                                         Goddard A,
                                                                                                                                                                                                                                                                                                                                               1999-551358/46.
                                                                                                                                                                                                                                                                                                                                                           P-PSDB; AAY41701
07 - MAY - 1998;
07 - MAY - 1998;
13 - MAY - 1998;
13 - MAY - 1998;
15 - MAY - 1998;
22 - MAY - 1998;
22 - MAY - 1998;
22 - MAY - 1998;
                                                                                                                                                                                                               -MAY-1998;
-MAY-1998;
-MAY-1998;
                                                                                                                                                                                                                                                  28-MAY-1998;
30-JUL-1998;
                                                                                                                                                                                                                                                                         11-SEP-1998
                                                                                                                                                                                                                                                                                                                         MI,
                                                                                                                                                                                                                                                                                                                                                                                                                               Claim
                                                                                                                                                                                                                                                                                                                         Wood
                                                                                                                                                                  18-
22-
22-
22-
22-
28-
```

; Gaps gtcccatccattaagccatcacttctggaagattaaagttgtcggacatggtgacagctg gtcccattccattaagccatcacttctggaagattaaagttgtcggacatggtgacagctg Length 4650; Indels 20; 18; DB Score 2033.8; Pred. No. 0; 1; Mismatches ï, 89.8%; Best Local Similarity 99.0 Matches 2066; Conservative Query Match 22 589 649 142 69 82 60, 202 g 8 염 à ò ò ò

```
1068
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ctctcagaggtagcaaaggaacatattgggaaggagggatccgggctgtaggctttgtgc 1401
                                                       441
                                                                                          501
                                                                                                                            561
                                                                                                                                                               621
                                                                                                                                                                                                  681
                                                                                                                                                                                                                                    741
                                                                                                                                                                                                                                                                     801
                                                                                                                                                                                                                                                                                                        861
                                                                                                                                                                                                                                                                                                                                          921
                                                                                                                                                                                                                                                                                                                                                                              981
                     381
                                                                                                                                                                                                                                   cacctactcttgacaagctcgctgccgaaggagttaaactggagaactactatgtccagc
                                                                                                                                                                                                                                                                               atgacttgtatgaaaacgacaatgctgcctgggactatgacaatggcatatactccacac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   agatgtacactcagagagtacagcaaatcttagcttcccataaccccacaaagcctatat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             gcattatcatttactcttcagataatggtggccagcctacggcaggaggagtaactggc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     agatgtacactcagagagtacagcaaatcttagcttcccataaccccacaaagcctatat
    gegeeegeageggggggggggggtggtteteegegtggagteteacetgggacetgagt
                                                                                                                             ctcaagctggagagaaactagagcccagcacaacttccacctcccagccccatctcattt
                                                                                                                                                                                                  cacctactcttgacaagctcgctgccgaaggagttaaactggagaactactatgtccagc
                                                                                                                                                                                                                                                                      ttcaacattctatcataagacctacccaacccaactgtttacctctggacaatgccaccc
                                                                                                                                                                                                                                                                                                         tacctcagaaactgaaggaggttggatattcaacgcatatggtcggaaaatggcacttgg
                                                                                                                                                                                                                                                                                                                                           gtttttacagaaaagaatgcatgcccaccagaagaggatttgataccttttttggttccc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ttttatatattgcctatcaagctgttcattcaccactgcaagctcctggcaggtatttcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          tagatgaagcaatcaacaacgtgacattggctctaaagacttatggtttctataacaaca
                                                         gtcctggaaagatgctagcaatggggggggctggcaggattctggatcctctgcctca
                                                                                          cttatggttacctgtcctggggccaggccttagaagaggaggaagaaggggccttactag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1222
                                                                                                                                                                                                                                                                                                                                                                                                                               1549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1282
                                                                                                                                                                                                                  1189
                                                                                                                                                                                                                                                    1249
                                                                                                                                                                                                                                                                                                                         1369
                                                                                                                                                                                                                                                                                                                                                            1429
                                                                                                                                                                                                                                                                                                                                                                                              1489
                                                                                                                                                                                                                                                                                                                                                                                                                                                  1042
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1849
   829
                     322
                                     889
                                                        382
                                                                        949
                                                                                          442
                                                                                                                           502
                                                                                                                                                               562
                                                                                                                                                                                                  622
                                                                                                                                                                                                                                                                      742
                                                                                                                                                                                                                                                                                                        802
                                                                                                                                                                                                                                                                                                                                          862
                                                                                                                                                                                                                                                                                                                                                                             922
                                                                                                                                                                                                                                                                                                                                                                                                                982
                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
qq
                                     g
                                                       ò
                                                                       Dp
                                                                                          ò
                                                                                                          d
                                                                                                                           ò
                                                                                                                                            q
                                                                                                                                                               δ
                                                                                                                                                                             g
                                                                                                                                                                                                 Q
                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                     οy
                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                              QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                     δ
```

```
1461
                                                                                                          1581
                                                                                                                                                1641
                                                                                                                                                                   2207
                                                                                                                                                                                      1701
                                                                                                                                                                                                         2267
                                                                                                                                                                                                                            1761
                                                                                                                                                                                                                                                                  1881
                                                                                                                                                                                                                                                                                                                            2447
                                                                                                                                                                                                                                                                                                                                                                  2507
                                                                                                                                                                                                                                                                                                                                                                                     2001
                                                                                                                                                                                                                                                                                                                                                                                                        2567
                                                                                                                                                                                                                                                                                                                                                                                                                        ggtaccccactctcatttcactggctgaaggacagattgatgaggacattcaactagatg 1521
                                                                                                                                                                                                                                                atcctggctacagcgactgggtccccctcagtctttcagcaacctgggaccgaaccggt 2327
                                                                                                                                                                                                                                                                                                                                               tacggaggctctcacagttcaacaaaactgcagtgccggtcaggtatccccccaaagacc 1941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               nootropic; neuroprotective; antibacterial;
                                                                                                                                               atcctggctacagcgactgggtccccctcagtctttcagcaacctgggaccgaaccggt
                                                                             gctatgatatctgggagaccataagtgagggtcttcgctcaccccgagtagatattttgc
                                                                                                                   tctggaacactgcaatccagtcagccatcagagtgcagcactggaaattgcttacaggaa
                                                                                                                                                                                                                                                                                                                             ccgacccatatgagagggtggacctatctaacaggtatccaggaatcgtgaagaagctcc
                                                                                                                                                                                                                                                                                                                                                         ccagaagtaaccctaggctcaatggaggggtctagggaccatggtatagagagaaacca
                             atagcccacttctgaaaaaaaaagggaacagtgtgtaaggaacctgtgcacatcactgact
                                                                                                                                                                                                                                                                                                          ccgacccatatgagaggtggacctatctaacaggtatccaggaatcgtgaagaagctcc
                                                                                                                                                                                                                                                                                                                                                                                                        ccagaagtaaccctaggctcaatggaggggtctggggaccatggtataaagaggaaacca
                                                                                                                                                                                                                                                                                                                   antiinflammatory; vasotropic; antitumor; gene therapy; human; ss
                                                                                                                                                                                                                                                                                                                                                                                                                                                               agaagaaacagcagaaagcagtctcaggttcaacttgccattcaggt 2107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2628 agaagaaacagcagaaagcagtctcaggtaaaccagcaaatttggct 2674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human 23553 sulfatase polypeptide coding sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /*tag= a
/product= "23553 sulfatase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
1..1800
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  standard; cDNA; 1800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             23553 sulfatase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200155411-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        25-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sulfatase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAH46867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAH46867;
                                                                                                                                                                                                                                                                                    2328
                                                                                                                                                                                                                                                                                                                          2388
                                                                   1462
                                                                                    2029
                                                                                                                                                1582
                                                                                                                                                                                    1642
                                                                                                                                                                                                        2208
                                                                                                                                                                                                                                              2268
                                                                                                                                                                                                                                                                                                                                                                 2448
                                                                                                                                                                                                                                                                                                                                                                                    1942
         1909
                            1402
                                                                                                         1522
                                                                                                                           2089
                                                                                                                                                                  2149
                                                                                                                                                                                                                          1702
                                                                                                                                                                                                                                                                  1762
                                                                                                                                                                                                                                                                                                        1822
                                                                                                                                                                                                                                                                                                                                               1882
                                                                                                                                                                                                                                                                                                                                                                                                       2508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                2061
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAH46867
         q
                                                                                                                                             ò
                                                                                                                                                                g
                                                                                                                                                                                                    С
                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                g
                                                                  ò
                                                                                    g
                                                                                                        ò
                                                                                                                        qq
                                                                                                                                                                                    ö
                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
1;
                                                                                                                                                                                                                                                                                        The invention provides 22438, 23553, 25278 or 26212 human sulfatase polypeptides and polynucleotides. The sulfatase genes and polypeptides are useful for treating disorders involving the brain such as cerebrovascular diseases, infections such as acute meningitis, demyelinating diseases including multiple sclerosis, degenerative diseases sfecting the cerebral cortex including Alzheimer's disease and Pick disease, spinocerebellar degenerations including spinocerebellar ataxias including Friedreich ataxia, and ataxia telangiectasia, degenerative diseases affecting motor neurons including amyotrophic lateral sclerosis, inborn errors of metabolism such as leukodystrophies, toxic and acquired metabolic diseases, including vitamin deficiencies, and neurocutaneous syndromes (phakamatoses) including neurofibromotosis.
                                                                                                                                                                                                       diagnosing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               743
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        324 atggeteceaggggetgtgeggggeateegeeteegeetteteeaaaggeetgtgtetgt 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                       Novel human sulfatase polypeptides useful for treating and diagnosir sulfatase-related disorders such as cerebrovascular diseases, acute
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  cctactcttgacaagctcgctgccgaaggagttaaactggagaactactattgtccagcct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 atggeteceaggggetgtgegggggeatecgeettectecacagggetgtetgt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    cctggaaagatgctagcaatggggggggcgctggcaggattctggatcctctctcact
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                atcctagcggatgatcagggatttagagatgtgggttaccacggatctgagatttaaaaca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        tatggttacctgtcctggggccaggccttagaagaagaagaagaaggggccttactagct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           atcctagcggatgatcagggatttagagatgtgggttaccacggatctgagattaaaaaca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               atttgcacaccatccaggagtcagtttattactggaaagtatcagatacacaccggactt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           atttgcacaccatccaggagtcagtttattactggaaagtatcagatacacaccggactt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          caacattctatcataagacctacccaacccaactgtttacctctggacaatgccacccta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 1800;
                                                                                                                                                                                                                                   meningitis, multiple sclerosis, degenerative diseases and tumor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                <u>т</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 1800 BP; 534 A; 434 C; 423 G; 409 T; 0 other;
                                                                                                                                  Tsai
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4,
                                                                                                                                Rudolph-Owen LA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 1781.6;
Pred. No. 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mismatches
                                                                                                                                                                                                                                                              Claim 2; Fig 15; 180pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ó
                                                                                                                                Williamson M,
                                                                                                   (MILL-) MILLENNIUM PHARM INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             78.6%;
ilarity 99.7%;
Conservative
                                         31-JAN-2001; 2001WO-US03266
                                                                     2000US-0495823
                                                                                                                                                              2001-476214/51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 1795; Conserv
                                                                                                                                                                           P-PSDB; AAB85483.
                                                                                                                                Glucksmann MA,
                                                                     31-JAN-2000;
             02-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       744
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  421
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

σ

```
1163
                                                                                                                                                                                                                                                  1043
                                                                                                                                          1103
                                                                                                                                                                                                                                                                                     attatcatttactcttcagataatggtggccagcctacggcaggaggagtaactggcct 1343
                                                                                                                                                                                                                                                                                                       taccccactctcatttcactggctgaaggacagattgatgaggacattcaactagatggc 1200
                                                                                                                                                                                                                                                                                                                                                                                                                                tatgatatctgggagaccataagtgagggtcttcgctcaccccgagtagatattttgcat 1583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    aacattgaccccatatacacccaaggcaaaaatggctcctggggcagcaggctatgggatc 1643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       tggaacactgcaatccagtcagccatcagagtgcagcactggaaattgcttacaggaaat 1703
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          cotggctacagogactgggtccccctcagtcttcagcaacctgggaccgaaccggtgg 1763
                                                                                                                                                                                                                cactaccgatccattatcaacataaacaggaggagatatgctgccatgcttcctgctta 1223
                                                                                                                                                                                                840
                                          ttgggaagtggggattactatacacactacaaatgtgacagtcctgggatgtgtgggctat
                                                                                                        gacttgtatgaaaacgacaatgctgcctgggactatgacaatggcatatactccacacag
                                                                                                                          gacttgtatgaaaacgacaatgctgcctgggactatgacaatggcatatactccacacag
                                                                                                                                                   taccccactctcatttcactggctgaaggacagattgatgaggacattcaactagatggc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               atgtacactcagagagtacagcaaatcttagcttcccataaccccacaaagcctatattt
                                                                                                                                                                             ttatatattgcctatcaagctgttcattcaccactgcaagctcctggcaggtatttcgaa
cctcagaaactgaaggaggttggatattcaacgcatatggtcggaaaatggcacttgggt
                                  ttttacagaaaagaatgcatgcccaccagaagaggatttgataccttttttggttccctt
                                                                                       ttgggaagtggggattactatacacactacaaatgtgacagtcctgggatgtgtgggctat
                                                                                                                                                                                                                                                                                                                                                                                                      1021
                                                                                                                                                                                                                                                                                                                                                                           1081
                                                                                                                                                                                                                                                                                                                                                                                                              1141
                                                                                                                                                                                                                                                                                                                                                                                                                               1524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1764
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1441
                481
                                                  541
                                                                                                                                          1044
                                                                                                                                                                            1104
                                                                                                                                                                                                                1164
                                                                                                                                                                                                                                                 1224
                                                                                                                                                                                                                                                                                     1284
                                                                                                                                                                                                                                                                                                                       1344
                                                                                                                                                                                                                                                                                                                                                           1404
                                                                                                                                                                                                                                                                                                                                                                                             1464
                                                                                                                                                                                                                                                                                                                                                                                                                                                 1201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1501
804
                                 864
                                                                     924
                                                                                     601
                                                                                                                         661
                                                                                                                                                           721
                                                                                                                                                                                                                                 841
                                                                                                                                                                                                                                                                    901
                                                                                                                                                                                                                                                                                                     196
                                                                                                       984
                                                                                                                                                                                              781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                              οŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
               셤
                                                 g
                                                                                     g
                                                                                                                    g
                                                                                                                                          ò
                                                                                                                                                        8
                                                                                                                                                                                           a
                                                                                                                                                                                                               à
                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                      ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
ò
                                  ò
                                                                     ò
                                                                                                       ò
                                                                                                                                                                           ò
```

agaagtaaccctaggctcaatggagggtctggggaccatggtataaagaggaaaccaag agaagtaaccctaggctcaatggaggggtctagggaccatggtatagagaggaaaccaag aaaaaqaaqccaaqcaaaaatcaqqctqaqaaaaaqcaaaaqqaaaqc-aaaaaaqaaq aagaaaacagcagaaagcagtctcaggttcaacttgccattcaggtgttacttgtggataa 1561 1944 1621 2004 1681 2063 1884 ŏ q q δ 9 δ g

9 26 RESULT AAS227

AAS22726 standard; cDNA; 1375

ВР

AAS22726;

(first entry) 24-OCT-2001

Human cDNA encoding a novel human protein #292.

Human; novel protein; ss; Antianaemic; osteopathic; antiinflammatory; immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic; anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral; antibocterial; antiallergic; describing antibocterial; antiallergic; describing antibocterial; antiallergic; describing antibocterial; immunogen; antibody; gene therapy; neurological disorder; thrombolytic; immunogen; antibody; gene therapy; neurological disorder; barkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis; tissue regeneration; immune disorder. 

Homo sapiens

WO200155437-A2.

02-AUG-2001.

25-JAN-2001; 2001WO-US02623

25-JAN-2000; 2000US-0491404

(HYSE-) HYSEQ INC

RT; Drmanac Tang YT, Liu C,

WPI; 2001-451939/48. P-PSDB; AAU14421 diseases for treating anti-inflammatory diseas for regenerating bone and cartilage useful for system disorders, and polypeptides Isolated nervous

Claim 1; Page 699; 894pp; English

proteins or their active domains. The polypeptides, polynucleotides and proteins or their active domains. The polypeptides, polynucleotides and antibodies raised against the polypeptides are used in a method of for treatment of a mammal and prevention of disorders caused by the aberrant protein expression or activity. The polypeptides can be used as molecular weight markers, food supplements, and in antibody production. The polypeptides are used to identify compounds which bind to the polypeptides. Polynucleotides of the invention are used as probes and primers, for sequencing, for chromosome or gene mapping, in the production of recombinant proteins, and in generating anti-sense DNA or RNA and in gene therapy. Polypeptides of the invention can be used to target drugs to a tumour, in assays to determine biological activity, to target antibodies/elicit an immune response, to determine duantitative protein levels, as tissue markers, and to isolate receptors or ligands. Polypeptides of the invention may also be useful in treating platelet disorders, stem cell disorders, regenerating bone, cartilage, tendon,

```
ö
ligament and/or nerve tissue, wound healing, treating burns, promoting the proliferation, differentiation and survival of stem cells, as a contraceptive, treating osteoporosis and osteoarthritis, anaemia. Alzheimer's, Parkinson's and Huntington's diseases, amylotrophic lateral sclerosis, stroke, immune deficiencies resulting from bacterial, viral or fungal infection or from autoimmunity, cancer, allergy, asthma, graft-versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory disease, nervous system disorders, and infection. The present sequence encodes a protein of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1023
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                atatactccacacagatgtacactcagagagtacagcaaatcttagcttcccataacccc 1088
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 atatactccacacagatgtacactcagagagtacagcaaatcttagcttcccataacccc 1203
                                                                                                                                                                                                                               903
                                                                                                                                                                                                                                                                                                  ggggccttactagctcaagctggagagagaaactagagcccagcacaacttccacctcccag 548
                                                                                                                                                                                                                                                                                                                                                                                             ccccatctcattttcatcctagcggatgatcaggggatttagaggatgtgggttaccacgga 608
                                                                                                                                                                                                                                                                                                                                                                                                                                                 668
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   tactatgtccagcctatttgcacaccatccaggagtcagtttattactggaaagtatcag 728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            848
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              aaatggcacttgggtttttacagaaaagaatgcatgcccaccagaagaggatttgatacc 908
                                                                                                                                                                                                                                                                                    ctctgcctcctcacttatggttacctgtcctgggggccaggccttagaagaggaggaagaa 488
                                                                                                                                                                                                                                                                                                                                                                  ggggccttactagctcaagctggagagagaactagagcccagcacaacttccacctcccag 663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gacaatgccaccctacctcagaaactgaaggaggttggatattcaacgcatatggtcgga 963
                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                           tctgagattaaaacacctactcttgacaagctcgctgccgaaggagttaaactggagaac
                                                                                                                                                                                                                                                                                                                                                                                                                                                             gacaatgccaccctacctcagaaactgaaggaggttggatattcaacgcatatggtcgga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             tactatgtccagcctatttgcacaccatccaggagtcagtttattactggaaagtatcag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        atacaccogacttcaacattctatcataagacctacccaaccgactgtttacctctg
                                                                                                                                                                                                       ö
                                                                                                                                                                             Length 1375;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                    Sequence 1375 BP; 399 A; 336 C; 360 G; 280 T; 0 other;
                                                                                                                                                                                                       ö
                                                                                                                                                                          Score 890; DB 22;
Pred. No. 5e-245;
); Mismatches 0;
                                                                                                                                                                      39.3%; Scor
100.0%; Pre
                                                                                                                                                                                        Best Local Similarity Luv.
Matches 890; Conservative
                                                                                                                                                                           Query Match
                                                                                                                                                                                                                               369
                                                                                                                                                                                                                                                                                                             544
                                                                                                                                                                                                                                                                                                                                                                                            549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1084
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1144
                                                                                                                                                                                                                                                                                   429
                                                                                                                                                                                                                                                                                                                                        489
                                                                                                                                                                                                                                                                                                                                                                                                                     664
                                                                                                                                                                                                                                                                                                                                                                                                                                               609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    904
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       964
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       696
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1089
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1264
  8888888888888
                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
```

```
immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic; anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral; antibacterial; antiallergic; dermatological; haemostatic; antiaschmatic; thrombolytic; immunogen; antibody; gene therapy; neurological disorder; Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis; tissue regeneration; immune disorder.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        primers, for sequencing, for chromosome or gene mapping, in the production of recombinant proteins, and in generating anti-sense DNA or RNA and in gene therapy. Polypeptides of the invention can be used to target drugs to a tumour, in assays to determine biological activity, to raise antibodies/elicit an immune response, to determine quantitative protein levels, as tissue markers, and to isolate receptors or ligands. Polypeptides of the invention may also be useful in treating platelet disorders, stem cell disorders, regenerating bone, cartilage, tendon, ligament and/or nerve tissue, wound healing, treating burns, promoting the proliferation, differentiation and survival of stem cells, as a contraceptive, treating osteoporosis and osteoarthritis, anamia, Alzheimer's, Parkinson's and Huntington's diseases, amplotrophic lateral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The invention relates to polynucleotides encoding novel human proteins or their active domains. The polypeptides, polynucleotides and antibodies raised against the polypeptides are used in a method of treatment of a mammal and prevention of disorders caused by the aberrant protein expression or activity. The polypeptides can be used as molecular weight markers, food supplements, and in antibody production. The polypeptides are used to identify compounds which bind to the polypeptides. Polynucleotides of the invention are used as probes and
                                                                                                                                                                                                                                                                                           protein; ss; Antianaemic; osteopathic; antiinflammatory;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               aptides useful for treating anti-inflammatory diseases, disorders, and for regenerating bone and cartilage -
Human cDNA encoding a novel human protein #56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 1; Page 256-257; 894pp; English.
                                                                                                                                      BP
                                                                                                                                    AAS22490 standard; cDNA; 1340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Isolated polypeptides useful
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25-JAN-2001; 2001WO-US02623.
                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Drmanac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2001-451939/48.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (HYSE-) HYSEQ INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Tang YT, Liu C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            P-PSDB; AAU14185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200155437-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   nervous system
                                                                                                                                                                                                                                                                                           novel
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25-JAN-2000;
                                                                                                                                                                                                               24-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             02-AUG-2001
                                                                                                                                                                           AAS22490;
                                                                                                                                                                                                                                                                                           Human;
                                                                                                                  AAS22490
                                                                                                  RESULT
                                      QQ
    ò
```

sclerosis, stroke, immune deficiencies resulting from bacterial, viral or fungal infection or from autoimmunity, cancer, albergy, asthma, graft-versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory disease, nervous system disorders, and infection.

present sequence encodes a protein of the invention Sequence 1340 BP; 374 A; 338 C; 369 G; 258 T; 1 other;

The

```
3;
                                                                                                                                                                                                                                                                                                                                                catatactccacacagatgtacactcagagagtacagcaaatcttagcttcccataaccc 1087
                                                                                                                                                                                                                                                                                                                                                                                                                               amatggcacttgggtttttacagaaaagaatgcatgcccaccagaagaggatttgatacc 1023
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ctggcaggtatttcgaacact--accgatccattatcaacataaacaggaggagatatgc 1204
                                                   543
                                                                  723
                                                                                                                                                                                                                            843
                                                                                                                                                                                                                                             788
                                                                                                                                                                                                                                                              903
                                                                                                                                                                                                                                                                                                                 aaatggcacttgggtttttacagaaaagaatgcatgccaccagaagaggatttgatacc 908
                  Gaps
                                                                                                               caggectgtgtctgtcctggaaagatgctagcaatgggggggggcgctggcaggattctggatc
                                                   tctgagattaaaacacctactcttgacaagctcgctgccgaaggagttaaactggagaac
                                                                                                                                                                                                           tactatgtccagcctatttgcacaccatccaggagtcagtttattactggaaagtatcag
                                                                                                                                                                                                                                             atacacaccggacttcaacattctatcataagacctacccaacccaactgtttacctctg
                                                                                                                                                                                                                                                                              gacaatgccaccctacctcagaaactgaaggaggttggatattcaacgcatatggtcgga
                                                                                                                                                                                                                                                                                                                                                                                   tgggatgtgtgtggctatgacttgtatgaaaacgacaatgctgcctgggactatgacaatgg
                                                                                                      ggggccttactagctcaagctggagagaaactagagcccagcacaacttccacctcccag
                                                                                                                                        ccccatctcattttcatcctagcggatgatcaggggatttagagatgtgggttaccacgga
  Length 1340;
                  Indels
 22;
                  44;
 Score 736.2; DB :
Pred. No. 8e-201; ; Mismatches 44
                 ;
0
32.5%;
ilarity 94.3%;
Conservative
         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          tgcc 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ggcc 1327
         Best Local Sim.
Matches 796;
 Query Match
                                                                                                                                                                                                                                                                                                                                                                  1024
                                                                                                                                                                                                                                                                                                                                                                                                    1084
                                                                                                                                                                                                                                                                                                                                                                                                                                     1144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1324
                                  369
                                                 484
                                                                  429
                                                                                    544
                                                                                                                                                                                                                                                             844
                                                                                                                                                                                                                                                                                               904
                                                                                                                                                                                                                                                                                                                                964
                                                                                                                                                                                                                                                                                                                                                                                   968
                                                                                                     489
                                                                                                                     604
                                                                                                                                      549
                                                                                                                                                      664
                                                                                                                                                                       609
                                                                                                                                                                                        724
                                                                                                                                                                                                          699
                                                                                                                                                                                                                          784
                                                                                                                                                                                                                                            729
                                                                                                                                                                                                                                                                              789
                                                                                                                                                                                                                                                                                                                                                 606
                                                                                                                                                                                                                                                                                                                                                                                                                     1028
                                                                                                                                                                                                                                                                                                                                                                                                                                                       1088
                                                a
                                                                                   a
                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                  ò
                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                   ò
                                                                                                    ò
                                                                                                                   셤
                                                                                                                                      ò
                                                                                                                                                    셤
                                                                                                                                                                       ò
                                                                                                                                                                                        8
                                                                                                                                                                                                         ò
```

```
libraries constructed from human colon cancer cell lines. The present invention also describes a method of detecting differentially expressed genes correlated with a cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample derived from a cell suspected of being cancerous, where detection of the differentially expressed gene product is correlated with a cancerous state of the cell from which the test sample was derived. The polynucleotides sequences can be used in a method for detecting mammalian cell. The polynucleotides can also be used as probes for detecting and mapping related genes. They can be used a hidgmosis and prognosis of diseases and disorders (e.g. identification of
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gatatctgggagaccataagtgagggtcttcgctcaccccgagtagatattttgcataac 1586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         mammalian
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           pre-metastatic or metastatic cancerous states, stages of cancer, or responsiveness of cancer to therapy). This is particularly for breast cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gatatotgggagaccataagtgagggtottcgctcaccccgagtagatattttgcataac 188
                                                                                                                                                                                                                                                                                                                                                                                                                      Sudduth-Klinger J;
               colon cancer cell line polynucleotide sequence SEQ ID NO:2335
                                             Human; colon cancer; tumour; diagnosis; gene expression product; probe; detection; cancerous state; metastasis; identification; breast cancer; oestrogen receptor-positive breast cancer; therapy; oestrogen receptor-negative breast cancer; lung cancer; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
                                                                                                                                                                                                                                                                                                                                                                                                                  J, Innis MA, Garcia PD, Sudduth-Klinge
Randazzo F, Kennedy GC, Pot D, Kassam
Crkvenjakov R, Dickson M, Drmanac S, I
Garcia V, Jones LW, Stache-Crain B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            27.6%; Score 625.2; DB 21; Length 729;
larity 97.3%; Pred. No. 4.4e-169;
Conservative 0; Mismatches 17; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polynucleotide library used to determine cancerous states of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               breast cancer, lung cancer, and colon cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BP; 232 A; 178 C; 163 G; 143 T; 13 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 1; Page 924-925; 1097pp; English.
                                                                                                                                                                                                                                                              98US-0085426.
98US-0085537.
98US-0085696.
98US-0105234.
                                                                                                                                                                                                                                 99WO-US10602
                                                                                                                                                                                                                                                                                                                                                                                                                      Escopedo
                                                                                                                                                                                                                                                                                                                                                                                                                                  Reinhard C, Giese K,
Lamson G, Drmanac R,
Leshkowitz D, Kita D,
                                                                                                                                                                                                                                                                                                                                                                   CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2000-126369/11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
Matches 644; Conserv
                                                                                                                                                                                                                                                                                                                                                                                   HYSEQ INC
                                                                                                                                                                                                                                                                                                                                                                   (CHIR ) CHIRON
                                                                                                                                                                                                                                                                                                                                                                                                                      Williams LT,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 729
                                                                                                                                                                                                                                                                                               15-MAY-1998;
21-OCT-1998;
27-OCT-1998;
                                                                                                                                 Homo sapiens
                                                                                                                                                                W09958675-A2
                                                                                                                                                                                                                                 13-MAY-1999;
                                                                                                                                                                                                                                                                14-MAY-1998;
15-MAY-1998;
                                                                                                                                                                                               18-NOV-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               negative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                   (HYSE-)
                 Human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
```

;;

BP.

CDNA; 729

AAA02344 standard;

œ

AAA02344

AAA02344;

XX

19-MAY-2000 (first entry)

```
1886
                                                                                                                                                                         1946
                                                                                                                                                                                                           2006
                                                                                                                                                                                                                                             2066
attgaccccatatacaccaaggcaaaaatggctcctgggcagcaggctatgggatctgg 1646
                                                                    ggctacagcgactgggtccccctcagtcttcagcaacctgggaccgaaccggtggcac 1766
                                                                                                                                                                                                                                                                                                                                                                                                            Sulfatase; 25278 sulfatase; nootropic; neuroprotective; antibacterial; antiinflammatory; vasotropic; antitumor; gene therapy; human; ss.
                                                                                                                                                                                          548
                                           ccatatgagagggtggacctatctaacaggtatccaggaatcgtgaagaagctcctacgg
                                                                                                                                                                         aggctctcacagttcaacaaaactgcagtgccggtcaggtaccccccaaagaccccaga
                                                                                                                                                                                                           aacactgcaatccagtcagccatcagagtgcagcactggaaattgcttacaggaaatcct
                                                                             ccatatgagagaggtggacctatctaacaggtatccaggaatcgtgaagaagctcctacgg
                                                                                                                                                                                           aggottotcacagttcaacaaaactgcagtgccggtcaggtatccccccaaagaccccaga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ъ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Tsai
                                                                                                                                                                                                                                                                                                                                                                                           Human 25278 sulfatase polypeptide encoding cDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ĽĄ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                334..2043
/*tag= a
/product= "25278 sulfatase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Rudolph-Owen
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                         BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Glucksmann MA, Williamson M,
                                                                                                                                                                                                                                                                                                                                         AAH46862 standard; cDNA; 2940
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (MILL-) MILLENNIUM PHARM INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    31-JAN-2000; 2000US-0495823.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31-JAN-2001; 2001WO-US03266
                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 2001-476214/51.
P-PSDB; AAB85482.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200155411-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                       Homo sapiens
                                                                                                                                                                                                                                                                              aa 2068
                                                                                                                                                                                                                                                                                               699
                                                                                                                                                                                                                                                                                                                                                                           25-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   02-AUG-2001
                                                                                                                                                                                                                                                                                                                                                          AAH46862
                                                                                                                                                                                                                                                                                         - a
                                 1647
                                                                                                                                                                                                           1947
                                                                                                                                                                                                                                                                              2067
                189
                                                 249
                                                                   1707
                                                                                                     1767
                                                                                                                      369
                                                                                                                                        1827
                                                                                                                                                       429
                                                                                                                                                                        1887
                                                                                                                                                                                          489
                                                                                                                                                                                                                            549
                                                                                                                                                                                                                                             2007
                                                                                                                                                                                                                                                                                               668
1587
                                                                                    309
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Key
                                                                                                                                                                                                                                                                                                                                 AAH46862
                                                                                                                                                                                       q
                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                  õ
                g
                                 ö
                                                 g
                                                                   ð
                                                                                   g
                                                                                                     ð
                                                                                                                   q
                                                                                                                                                      a
                                                                                                                                                                                                          ò
                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                              ò
                                                                                                                                       ð
                                                                                                                                                                        ò
```

```
The invention provides 22438, 23553, 25278 or 26212 human sulfatase polypeptides and polynucleotides. The sulfatase genes and polypeptides are useful for treating disorders livoviving the brain such as degenerative described in the cerebrovascular diseases, infections such as acute meningtis, demyelinating diseases including multiple sclerosis, degenerative diseases affecting the cerebral cortex including Alzhehumer's disease attaxias including spinocerebellar and pick disease, spinocerebellar degenerations including spinocerebellar degenerative diseases affecting motor neurons including amyotrophic lateral sclerosis, inborn errors of metabolism such as leukodystrophies, toxic and acquired metabolic diseases, including vitamin deficiencies, and neurocutaneous syndromes (phakamatoses) including neurofibromotosis. The present sequence represents a human 25278 sulfatase polypeptide encoding cDNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            441
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                  acute
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         tggcgatgcacaccctcactggcttctctctggtcagcctgctcagcttcggctacctgt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cctgggactgggcc----aagccgagcttcgtggccgacgggcccggggaggctggc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        tagcaatgggggggctggcaggattctggatcctctgcctcacttatggttacctgt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  agctcgctgccgaaggagttaaactggagaactactatgtccagcctatttgcacaccat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             cctgggggccaggccttagaagaagaagaagaagaagggccttactagctcaagctggagaga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                aactagagcccagcacaacttccacctcccagccccatctcatttcatcctagcggatg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   atcagggatttagagatgtgggttaccacggatctgagattaaaacacctacttgaca
Novel human sulfatase polypeptides useful for treating and diagn sulfatase-related disorders such as cerebrovascular diseases, ac meningitis, multiple sclerosis, degenerative diseases and tumor
                                                                                                                                                                                                                                                                                                                                                                                                                                                        15;
                                                                                                                                                                                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                 Sequence 2940 BP; 567 A; 915 C; 902 G; 556 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                   Score 596.2; DB 22;
Pred. No. 2.2e-160;
); Mismatches 598;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                   2; Fig 10; 180pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                     26.3%;
62.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Simi
Matches 1010;
                                                                    Claim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 577
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          620
 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
```

gagtacagcaaatcttagcttcccataaccccacaaagcctatattttatatattgcct 1116

1057

ŏ

acgacaatgctgcctgggactatgacaatggcatatactccacacagatgtacactcaga 1056

816

918 799 936 859 966

tecgeceacageageceaactgeetgeectgaeceaggtgaeactgeeacagaagetge 739

taagacctacccaacccaactgtttacctctggacaatgccaccctacctcagaaactga

757 680 817 740 877 800 937 860 266

ολ g ŏ g ò g ò g õ

aggaggttggatattcaacgcatatggtcggaaaatggcacttgggtttttacagaaaag

attactatacacactacaaatgtgacagtcctgggatgtgtgggctatgacttgtatgaaa

aatgcatgcccaccagaagagtttgatacctttttggttcccttttgggaagtgggg

(first

13

```
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                          1636
                                                                                                                                                                                                                                                        gegecagecatatectggecagecacageceteagegteeeetetteetetatgtggeet 1036
                                                                                                                  ccagtgacaatggtggccagactttctcggggggcagcaactggccgctccgaggacgca 1276
                                                                                                                                                                                                                            agaccataagtgagggtcttcgctcaccccgagtagatattttgcataacattgacccca 1596
                                                                                                                                                                                                                                                                                                    cggccatcagcgagggccgggcctcaccacgcacggagatcctgcacaaacattgacccac 1516
                                                                                                                                                                                                                                                                                                                     tatacaccaaggcaaaaaatggctcctgggcagcaggctatgggatctggaacactgcaa 1656
                                                                                                                                                                                                                                                                                                                                                                                                              attggatcccaccgcagacactggccaccttcccgggtagctggtggaacctggaacgaa 1696
                                                                                                                                                                                                                                                                                                                                                                                                                                tcacctcgtcaactggcaaaagtgtatggcttttcaacatcacagccgacccatatgaga 1836
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     999aagaacetggetggecageggeetgatgtggteegeaeeetgetggetegeetggeeg 1813
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       agttcaacaaaactgcagtgccggtcaggtatccccccaaagacccccagaagtaacccta 1956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gggtggacctatctaacaggtatccaggaatcgtgaagaagctcctacggaggctctcac
                                  ttatcaacataaacaggaggagatatgctgccatgcttcctgcttagatgaagcaatca
                                                                      acaacgtgacattggctctaaagacttatggtttctataacaacagcattatcatttact
                                                                                                                                   cttcagataatggtggccagcctacggcaggaggagtaactggcctctcagaggtagca
                                                                                                                                                                                         agggcacttattgggaaggtggcgtgcggggcctaggctttgtccacagtcccctgctca
                                                                                                                                                                                                           actgggtccccctcagtctttcagcaacctgggaccgaaccggtggcacaatgaacgga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ggctcaatggagggtctaggggaccatggtatagaggaaaccaagaaaaaagaagccaa
                                                                                                                                                                        aaggaacatattgggaaggagggatccgggctgtaggctttgtgcatagcccacttctga
                                                                                                                                                                                                                                               tttcactggctgaaggacagattgatgaggacattcaactagatggctatgatatctggg
                                                                                                                                                                                                                                                                                                                                                        tecagteagecateagagtgeageactggaaattgettaeaggaaateetggetaeageg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        aatataaccgcacagccatcccggtacgctacccagctgagaacccccggggctcatcctg
                        atcaagctgttcattcaccactgcaagctcctggcaggtatttcgaacactaccgatcca
                                                                                                                                                                                                                                                                                            =
=
=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1936
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1874
                       1117
                                         1037
                                                            1177
                                                                                                                                                    1217
                                                                                                                                                                                       1277
                                                                                                                                                                                                          1417
                                                                                                                                                                                                                            1337
                                                                                                                                                                                                                                              1477
                                                                                                                                                                                                                                                                1397
                                                                                                                                                                                                                                                                                  1537
                                                                                                                                                                                                                                                                                                                     1597
                                                                                                                                                                                                                                                                                                                                      1517
                                                                                                                                                                                                                                                                                                                                                                         1577
                                                                                                                                                                                                                                                                                                                                                                                                             1637
                                                                                                                                                                                                                                                                                                                                                                                                                                                1697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1754
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1957
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1934
     977
                                                                            1097
                                                                                                                1157
                                                                                                                                   1297
                                                                                                                                                                      1357
                                                                                                                                                                                                                                                                                                   1457
                                                                                                                                                                                                                                                                                                                                                        1657
                                                                                                                                                                                                                                                                                                                                                                                            1717
                                                                                                                                                                                                                                                                                                                                                                                                                               1777
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1837
                                                                            g
                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
     8
                                        g
                                                            ò
                                                                                                              a
                                                                                                                                                   Q
                                                                                                                                                                                     q
                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                       ò
                                                                                              ò
                                                                                                                                  δ
                                                                                                                                                                     ò
                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                              οχ
                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
The invention provides 22438, 23553, 25278 or 26212 human sulfatase polypeptides and polynucleotides. The sulfatase genes and polypeptides are useful for treating disorders involving the brain such as cerebrovascular diseases, infections such as acute meningitis, demyelinating diseases including multiple sclerosis, degenerative diseases affecting the cerebral cortex including Alzheimer's disease and Pick disease, spinocerebellar degenerations including spinocerebellar ataxias including Friedreich ataxia, and ataxia telangiectesia, degenerative diseases affecting motor neurons including amyotrophic tateral sclerosis, inborn errors of metabolism such as leukodystrophics, toxic and acquired metabolic diseases, including vitamin deficiencies, and neurocutaneous syndromes (phakamatoses) including neurofibromotosis.
                                              Sulfatase; 25278 sulfatase; nootropic; neuroprotective; antibacterial; antiinflammatory; vasotropic; antitumor; gene therapy; human; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           diagnosing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           atgggggcgctggcaggattctggatcctctgcctcctcacttatggttacctgtcctgg 461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        acute
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 cageceteggeegeteegeeceageeteeceacateatetteateeteaeggaegaa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ggatttagagatgtgggttaccacggatctgagattaaaacacctactcttgacaagctc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         gactgggcc----aagccgagcttcgtggccgacggggccgagggaggctggcga--g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               atgcacaccetcactggcttctctctggtcagcctgctcagcttcggctacctgtcctgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ggccaggccttagaagaggaagaagaaggggccttactagctcaagctggagaaacta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gagcccagcacaacttccacctcccagccccatctcattttcatcctagcggatgatcag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Novel human sulfatase polypeptides useful for treating and diagn sulfatase-related disorders such as cerebrovascular diseases, acmeningitis, multiple sclerosis, degenerative diseases and tumor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 1710;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               굕,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Seguence 1710 BP; 340 A; 554 C; 500 G; 316 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Tsai
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 22;
sulfatase polypeptide coding sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             296;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 594.4; DB 22
Pred. No. 5.2e-160;
); Mismatches 596;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Rudolph-Owen LA,
                                                                                                                                                                                                                                                sulfatase"
                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                                                                                                "25278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Claim 2; Fig 10; 180pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Glucksmann MA, Williamson M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (MILL-) MILLENNIUM PHARM INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            26.2%;
62.2%;
                                                                                                                                                                                                                                                                                                                                                                                                 31-JAN-2001; 2001WO-US03266.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  31-JAN-2000; 2000US-0495823.
                                                                                                                                                                                                                      /*tag= a
/product= '
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                            1..1710
/*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2001-476214/51.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          P-PSDB; AAB85482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       coding sequence
                                                                                                                                                                                                                                                                                                    WO200155411-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Simil
Matches 1007; C
                                                                                                                         sapiens
Human 25278
                                                                                                                                                                                                                                                                                                                                                  02-AUG-2001
                                                                                                                         Homo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δý
```

4

BP.

AAH46866 standard; cDNA; 1710

RESULT 10

AAH46866

AAH46866;

AXB

24 10:05:49 2002

Wed Jul

14

Page

1481 tcagccatcagaytgcagcactggaaattgcttacaggaaatcctggctacagcgactgg 1721 gctgttcattcaccactgcaagctcctggcaggtatttcgaacactaccgatccattatc 1181 aacataaacaggaggagatatgctgccatgctttcctgcttagatgaagcaatcaacaac 1241 acatattgggaaggagtccgggctgtaggctttgtgcatagcccacttctgaaaaac 1421 ctggctgaaggacagattgatgaggacattcaactagatggctatgatatctgggagacc 1541 ataagtgagggtcttcgctcaccccgagtagatattttgcataacattgaccccatatac 1601 accaaggcaaaaaatggctcctgggcagcaggctatgggatctggaacactgcaatccag 1661 getgccatccgcgtgggtgagtggaagctgctgacaggagaccccggctatggcgattgg 1308 tatacacactacaaatgigacagiccigggatgigggctaigaciigiatgaaaacgac 1001 cagcaaatcttagcttcccataaccccacaaagcctatatttttatatattgcctatcaa 1121 591 gacaatggtggccagactttctcggggggggacactggccgctccgaggacgcaagggc 948 821 881 atgcccaccagaagagatttgataccttttttggttcccttttgggaagtggggattac 941 701 761 351 gttggatattcaacgcatatggtcggaaaatggcacttgggtttttacagaaaagaatgc tacacctatgacaactgtgatggcccaggcgtgtgcggcttcgacctgcacgagggtgag aatgctgcctgggactatgacaatggcatatactccacacagatgtacactcagagagta gctgccgaaggagttaaactggagaactactatgtccagcctatttgcacaccatccagg agtcagtttattactggaaagtatcagatacacccggacttcaacattctatcataaga cctacccaacccaactgtttacctctggacaatgccaccctacctcagaaactgaaggag agccatatcctggccagccacagcctcagcgtcccctcttcctatgtggccttccag gataatggtggccagcctacggcaggaggagtaactggcctctcagaggtagcaaagga 1062 232 412 532 1002 1122 1242 829 949 1422 1009 1069 1542 542 762 942 592 709 1182 169 1302 889 1362 1482 1129 1602 1189 1249 702 292 352 822 882 649 1662 g 셤 g qq ö 셤 a ö g ò g ò g ò ò QQ ò qq ò g ò ò ò g ò ò g ò qq ò qq ò g ð g ò

```
Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase, ligase; hyperproliferative disorder; immunodeficiency disorder; autolimmune disorder; neurological disorder; metabolic disorder; inflammatory disorder; cardiovascular disorder; reproductive disorder; blood·related disorder; infectious disorder; gene therapy; cytostatic; anti arthritic; nephrotropic; anticoagulant; ss.
                                                                                               gacctatctaacaggtatccaggaatcgtgaagaagctcctacggaggctctcacagttc 1901
                                                                                                                                                                        aaccgcacagccatcccggtacgctacccagctgagaacccccgggctcatcctgacttt 1545
                                                                                                                                                aacaaaactgcagtgccggtcaggtatcccccaaagaccccagaagtaaccctaggctc 1961
gicccccicagicticagcaaccigggaccgaaccggiggcacaaigaacggaicacc 1781
                                                                                                                                                                                              2019
                                                                                                                                                                                                              atcccaccgcagacactggccaccttcccgggtagctggtggaacctggaacgaatggcc
                                                             gacctggctggccagcggcctgatgtggtccgcacctgctggctcgcctggccgaatat
                                               tcgtcaactggcaaaagtgtatggcttttcaacatcacagccgacccatatgagaggtg
                                                                                                                                                                                              1962 aatggaggggtctagggaccatggtatagaggaaaccaagaaaaagaagccaagca
                                                                                                                                                                                                                                                                                                                                                          cDNA encoding novel human enzyme polypeptide #421
                                                                                                                                                                                                                                                                                   BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2000US-0205515.
2000US-0209467.
2000US-0214886.
                                                                                                                                                                                                                                                                                  AAS41205 standard; cDNA; 629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000US-0179065.
2000US-0180628.
2000US-0184664.
2000US-0186350.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2000US-0224518.
2000US-0224519.
2000US-0225213.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2000US-0189874.
2000US-0190076.
2000US-0198123.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2000US-0216647,
2000US-0216880.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2000US-0217487
2000US-0217496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2000US-0218290.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000US-0220963
2000US-0220964
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000US-0225266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2000US-0225267
2000US-0225268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2000US-0225270
2000US-0225447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000US-0215135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0000US-0225214
                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200155301-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 16-MAR-2000;
17-MAR-2000;
18-APR-2000;
07-40N-2000;
07-40N-2000;
07-40L-2000;
07-40L-2000;
11-40L-2000;
11-40L-2000;
11-40L-2000;
26-40L-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            26-JUL-2000;
14-AUG-2000;
14-AUG-2000;
14-AUG-2000;
14-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              04-FEB-2000;
24-FEB-2000;
02-MAR-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            17-JAN-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         14-AUG-2000;
14-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            14-AUG-2000;
14-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14-AUG-2000;
                                                                                                                                                                                                                                                                                                                                  17-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    02-AUG-2001
                                                                                                                                                                                                                                                                                                         AAS41205;
                        1309
                                                1782
                                                                        1369
                                                                                               1842
                                                                                                                       1426
                                                                                                                                                1902
                                                                                                                                                                       1486
                                                                                                                                                                                                                       1546
                                                                                                                                                                                                                                                                     AAS41205
                                                                       g
                                                                                                                     Dp
                                                                                                                                                                       pp
                                                                                                                                                                                                                                                                                                qq
                                               οy
                                                                                               ò
                                                                                                                                                ŏ
                                                                                                                                                                                              ò
```

```
20000S - 022628
20000S - 0226681
20000S - 0226681
20000S - 0227182
20000S - 022934
20000S - 0229343
20000S - 0229345
20000S - 0229345
20000S - 0229345
20000S - 0229345
20000S - 0231243
20000S - 0231243
20000S - 0231243
20000S - 0231413
20000S - 0231413
20000S - 0231243
20000S - 0231240
20000S - 0231286
                                                                                                                                                                                                                                                                                                                                                                                               2000US-0237039.
2000US-0237040.
2000US-0239935.
2000US-0239937.
                                                                                                                                                                                                                                                                                                                                                                                                                              2000US-0240960.
2000US-0241221.
2000US-0241785.
2000US-0241786.
2000US-0241808.
2000US-0241809.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2000US-0241826.
2000US-0244617.
2000US-0246474.
2000US-0246475.
2000US-0246475.
                                                                                                                                                                                                                                                       2000US-0233064.
2000US-0233065.
2000US-0234223.
2000US-0234274.
                                                                                                                                                                                                                                                                                                               2000US-0235834.
2000US-0235836.
2000US-0236327.
                                                                                                                                                                                                                                                                                                                                                2000US-0236368
2000US-0236369
                                                                                                                                                                                                                                                                                                                                                               2000US-0236370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000US-0246532
                                                                                                                                                                                                                                                                                                                                        2000US-0236367
                                                                                                                                                                                                     14-SEP-2000;
14-SEP-2000;
14-SEP-2000;
14-SEP-2000;
14-SEP-2000;
14-SEP-2000;
14-SEP-2000;
                                                                                                                                                                                                                                                                              21-SEP-2000;
25-SEP-2000;
25-SEP-2000;
26-SEP-2000;
27-SEP-2000;
                                                                                                                                                                                                                                                                                                                                      29-SEP-2000;
29-SEP-2000;
29-SEP-2000;
29-SEP-2000;
02-OCT-2000;
02-OCT-2000;
02-OCT-2000;
02-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-NOV-2000;
08-NOV-2000;
08-NOV-2000;
                                                                                                                                                                      08-SEP-2000;
08-SEP-2000;
08-SEP-2000;
12-SEP-2000;
                              22-AIG-2000;
22-AUG-2000;
                                                                                                                                                                                                                                                               14-SEP-2000;
21-SEP-2000;
                                                                                                                                                                                                                                                                                                                       27-SEP-2000;
29-SEP-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        08-NOV-2000;
08-NOV-2000;
```

20000S - 0246610 2000US - 0246611 2000US - 0249203 2000US - 0249209 2000US - 0249210 2000US - 0249210 2000US - 0249212 2000US - 0249213 2000US - 0249213 2000US - 0249213 2000US - 0249214 2000US - 0249214 2000US - 0249214 2000US - 0249216 2000US - 0249216 2000US - 0249216 2000US - 0249216 2000US - 0249217 2000US - 0249218 2000US-0251479. 2000US-0251856. 2000US-0251868. 17-NOV-2000; 2 17-NOV-2000; 2 01-DEC-2000; 2 05-DEC-2000; 2 05-DEC-2000; 2 05-DEC-2000; 2 06-DEC-2000; 2 08-DEC-2000; 2 08-DEC-2000; 2 08-DEC-2000; 2 08-DEC-2000; 2 17 - NOV - 2000; 2 

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465566/50. P-PSDB; AAU23335.

Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases

Claim 4; SEQ ID No 431; 1180pp; English.

The present invention relates to the isolation of novel human enzyme polypeptides (AAU22915-AAU3814), and the CDNA and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the disorders including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g. arthritis), neurological disorders (e.g. AIzheimer's disorders (e.g. arthritis), neurological disorders (e.g. atherosclerosis), metabolic disorders (e.g. phenylkeronuria), inflammatory disorders (e.g. arthritis), neurological disorders (e.g. atherosclerosis), metabolic disorders (e.g. phenylkeronuria), inflammatory disorders (e.g. infertility) and infectious disorders (e.g. Influenza). The polynuclectides of the invention can also be used in gene therapy. AAS40785-AAS41684 represent cDNA sequences encoding for the novel human coryme polypeptides of the invention. at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 629 BP; 189 A; 151 C; 124 G; 162 T; 3 other;

```
tgacagtcctgggatgtgtggctatgacttgtatgaaaacgacaatgctgcctgggacta 1018
                                                                                                                                                                                                                                                                                     Drmanac RT;
                                                                                                                                                22-DEC-2000; 2000WO-US35017.
                                                                                                                                                                                  23-DEC-1999; 99US-0471275.
21-JAN-2000; 2000US-0488725.
25-APR-2000; 2000US-0552317.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 539; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            neurological disorders.
               neurological disorder;
Alzheimer's disease;
                                                                                                                                                                                                                                                                                                                      WPI; 2001-457603/49.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                      (HYSE-) HYSEQ INC
                                                                                                                                                                                                                                                                                     Tang YT, Liu C,
                                                                                                                                                                                                                                                                                                                                          P-PSDB; AAM25415
                                                                                  WO200153455-A2
                                                  Homo sapiens.
                                                                                                                   26-JUL-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 201
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Вp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                .;
7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       antific ammentory; antificened in the current interests of antific ammentory; antificened in the control, nument numentory antificened antificial antificial antificened antif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        aacgacaatgotgoctgggactatgacaatggcatatactccacacagatgtacactcag 1055
                                                                                577 atcagggatttagagatgtgggttaccacggatctgagattaaaaacacctactctgaca 636
                                                                                                                                                                                                                   ccaggagtcagtttattactggaaagtatcagatacacaccggacttcaacattctatca 756
                                                                                                                                                                                                                                    taagacctacccaacccaactgtttacctctggacaatgccaccctacctcagaaactga 816
                                                                                                                                                                                                                                                                                                                     taagacctacccaacccaactgtttacctctggacaatgccaccctacctcagaaactga 241
                                                                                                                                                                                                                                                                                                                                                    gattactatacacactacaaatgtgacagtcctgggatgtgtggctatgacttgtatgaa 995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          637 agctcgctgccgaaggagttaaactggagaactactatgtccagcctatttgcacaccat 696
                                                                                                               2 attoggcacgaggagatgtgggttaccacggatctgagattaaaacacctactcttgaca 61
                                                                                                                                                                Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
                                                5
             Length 629;
                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             protein encoding cDNA sequence SEQ ID NO:191.
               DB 22;
                             .6e-155
            Score 575.2; I
Pred. No. 9.6e.
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          attatcaacataaacaggagga 1197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           601 attatccacntaaccggangaa 622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAH99356 standard; cDNA; 562
             25.4%;
97.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16-OCT-2001 (first entry)
                                               Conservative
                               Similarity
                                               Matches 605;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAH99356;
                 Query Match
                                  Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        542
                                                                                                                                                                                                                                                                                                                   182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1056
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1176
                                                                                                                                                                                  62
                                                                                                                                                                                                                                                  122
                                                                                                                                                                                                                                                                                                                                                                                                                                                        302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         966
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          482
                                                                                                                                                                                                                                                                                     757
                                                                                                                                                                                                                                                                                                                                                                                                                      877
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       936
                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                               合
                                                                                                                                                ò
                                                                                                                                                                                  g
                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
```

```
AAM199166 to AAH99904 encode the human proteins given in AAW25225 to Calls they are expressed in, such as: antinities based on the tissues and cells they are expressed in, such as: antinities antinheumatic; antinarthritic; immunosuppressive; antibacterial; endocrine; cardiant; cardiarthritic; immunosuppressive; antibacterial; endocrine; cardiant; cardiant; or cardiovascular; antianemnic; antialgregant; hademostatic; uninterary; antidiabetic; osteopathic; dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic; antidiabetic; softens and polynucleotides are useful for screening for gonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, haematopoietic disorders, osteoporosis, severe combined immunodeficiency, eccama allergic hinitis, asthma, diabetes, cancer, multiple sclerosis, depression, allergic antidials, asthma, diabetes, cancer, multiple sclerosis, depression, allergic antidials, asthma, diabetes, neurodegenerative and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  778
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                928
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
Parkinson's disease; neurodegenerative disorder;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               779 tttacctctggacaatgccaccctacctcagaaactgaaggaggttggatattcaacgca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   tttacctctggacaatgccaccctacctcagaaactgaaggaggttggatattcaacgca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   atttgataccttttttggttccctttttgggaagtggggattactatacacactacaaatg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Isolated human polynucleotides encoding polypeptides, useful for treatment and diagnosis of e.g. cancer, ulcers and HIV infection % \left( 1\right) =\left\{ 1\right\} =
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 562 BP; 170 A; 132 C; 105 G; 155 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 538.4; DB 22;
Pred. No. 3.4e-144;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 1; Page 383; 1217pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23.8%;
```

ö

```
2000US-0237040.
2000US-0239935.
2000US-0239937.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2000US-0241221.
2000US-0241785.
2000US-0241786.
                                                                                                                                                                         2000US-0229343.
2000US-0229344.
2000US-0229345.
2000US-0229345.
                                                                                                                                                                                                                                                                           2000US-0229513.
2000US-0230437.
2000US-0230438.
2000US-0231242.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  2000US-0231414.
2000US-0232080.
2000US-0232081.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2000US-0234997.
2000US-0234998.
2000US-0235484.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2000US-0235834.
2000US-0235836.
2000US-0236327.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000US-0241787.
2000US-0241808.
2000US-0241809.
2000US-0241826.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000US-0246474.
2000US-0246475.
2000US-0246476.
2000US-0246477.
                              2000US-0226681.
2000US-0226868.
2000US-0227182.
2000US-0227009.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2000US-0231968.
2000US-0232397.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2000US-0234223.
2000US-0234274.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2000US-0236368
2000US-0236369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2000US-0236370
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20000S-0237038
20000S-0237039
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            :000US-0240960.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2000US-0246478
2000US-0246523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2000US-0246524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000US-0246525
2000US-0246526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2000US-0246532
2000US-0246609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2000US-0246610
2000US-0246611
                                                                                                                                                                                                                                                                                                                                                                                  2000US-0231243
                                                                                                                                                                                                                                                                                                                                                                                                    2000US-0231244
2000US-0231413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2000US-0232398
2000US-0232399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2000US-0232400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2000US-0232401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2000US-0233063
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2000US-0233064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2000US-0236802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0000S-0237037
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2000US-0244617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2000US-0246527
                                                                                                                           2000US-0228924
2000US-0229287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2000US-0236367
       2000US-0226279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12 - SEP - 2000; 2
14 - SEP - 2000; 2
15 - SEP - 2000; 2
25 - SEP - 2000; 2
26 - SEP - 2000; 2
26 - SEP - 2000; 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          29-SEP-2000; 29-SEP-2000; 29-SEP-2000; 20-OCT-2000; 20-OC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               02-0CT-2000; 2
13-0CT-2000; 2
13-0CT-2000; 2
20-0CT-2000; 2
20-0CT-2000; 2
20-0CT-2000; 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20-0CT-2000;
20-0CT-2000;
20-0CT-2000;
20-0CT-2000;
01-NOV-2000;
08-NOV-2000;
  18-AUG-2000; 22-AUG-2000; 22-AUG-2000; 22-AUG-2000; 23-AUG-2000; 23-AUG-2000; 20-SEP-2000; 20-SE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        27-SEP-2000;
27-SEP-2000;
29-SEP-2000;
29-SEP-2000;
       Human; respiratory antigen; respiratory disorder; throat disorder; lung disorder; note disorder; lung cancer; gene therapy; cytostatic; anti allergic; anti asthmatic; anti inflammatory; olfactory; respiratory active; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Genomic sequence #480 encoding for novel human respiratory antigen.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2000US-0179065.
2000US-018664.
2000US-018664.
2000US-018974.
2000US-019974.
2000US-0199123.
2000US-020515.
2000US-020515.
2000US-020515.
2000US-0216467.
2000US-021647.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAS28640 standard; DNA; 1615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20000S-0224518.
20000S-0224519.
20000S-0225213.
20000S-0225214.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2000us-0225266.
2000us-0225267.
2000us-0225268.
2000us-0225268.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2000US-0217496
2000US-0218290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2000US-0220963.
2000US-0220964.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2000US-0225758
2000US-0225759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        17-JAN-2001; 2001WO-US01333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO200155448-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    31-JAN-2000;
04-FEB-2000;
02-MAR-2000;
16-MAR-2000;
17-MAR-2000;
19-MAY-2000;
19-MAY-2000;
07-JUL-2000;
07-JUL-2000;
07-JUL-2000;
11-JUL-2000;
11-AUG-2000;
14-AUG-2000;
14-AUG-2000;
14-AUG-2000;
14-AUG-2000;
14-AUG-2000;
14-AUG-2000;
14-AUG-2000;
14-AUG-2000;
14-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 14-AUG-2000;
14-AUG-2000;
14-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   07-NOV-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          02-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAS28640;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 13
                                                                          1019
                                                                                                                                                                       1079
                                                                                                                                                                                                                                                                           1139
                                                                                                                                                                                                                                                                                                                            441
                                                                                                                                                                                                                                                                                                                                                                          1199
                                                                                                                                                                                                                                                                                                                                                                                                                           501
                         261
                                                                                                                           321
                                                                                                                                                                                                                        381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAS28640
                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A CONTRACTOR OF THE PROPERTY O
                                                                          ò
                                                                                                                      g
                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                        a
```

Page 18

```
1444 ctgtgcacatcactgactggtaccccactctcatttcactggctgaaggacagattgatg 1503
                                                                                                                                                                                                                             ggaaattgcttacaggaaatcctggctacagcgactgggtccccctcagtcttcagca 1743
                                                                                                                                                                                                                                                                                     ggcttttcaacatcacagccgacccatatgagaggtggacctatctaacaggtatccag 1863
                                                                                                                 cccgagtagatattttgcataacattgaccccatatacaccaaggcaaaaatggctcct 1623
                                                                                                                                                                        gggcagcaggctatgggatctggaacactgcaatccagtcagccatcagagtgcagcact 1683
                                                                                                                                                                                                                                                                                                                                                                                                      gaatcgtgaagaagctcctacggaggctctcacagttcaacaaaaactgcagtgccggtca 1923
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ggtatcccccaaagaccccagaagtaaccctaggctcaatggaggggtctagggaccat 1983
                                                                                                                                                                                        478 atgiggtccgcacctgctggctcgcctggccgaatataaccgcacagccatcccggtac 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         getacceagetgagaacceeegggeteateetgaetttaatgggggtgettggggggeeet 597
                                                                                                                                tgatgcacatcactgactggtacccgaccctggtgggtctggcaggtggtaccacctcag
                                                          aggacattcaactagatggctatgatatctgggagaccataagtgagggtcttcgctcac
                                                                                                                                                                                                                                              ggtatagagagaaaccaagaaaaagaagccaagca 2019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                598 gggccagtgatgaggaaggaggaaggaaggga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAA02370 standard; cDNA; 1082
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99WO-US10602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO9958675-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           L3-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19-MAY-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18-NOV-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       14-MAY-1998;
                             61
                                                                                                                 1564
                                                                                                                                            181
                                                                                                                                                                        1624
                                                                                                                                                                                                                               1684
                                                                                                                                                                                                                                                           301
                                                                                                                                                                                                                                                                                                                                               1804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1984
                                                                                                                                                                                                                                                                                                                                                                                                      1864
                                                                                                                                                                                                                                                                                                                                                                                                                                                             1924
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAA02370
                                                                                   g
                                                                                                                                         οp
                                                                                                                                                                        à
                                                                                                                                                                                                   qq
                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                 Ω
                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                          ф
                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Db
                          qq
                                                          ò
                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The present invention relates to the isolation of novel human respiratory antigens (AAU17685-AAU17975), and cDNA and genomic sequences encoding for these polypeptides. The sequences of the invention are useful for preventing, treating and/or prognosing disorders related to the respiratory system including throat disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis), lung disorders e.g. pneumonia, allergic disorders e.g. asthma, pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of the respiratory tissues e.g. lung cancer. The polyuucleotide sequences of the invention are useful in gene therapy and antisense therapy. AAS28161-AAS28764 represent genomic sequences encoding for novel Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Isolated polypeptide for treating, preventing and/ or prognosing disorders related to the respiratory system including respiratory cancers and also for testing and detection e.g. diagnosis -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 22; Length 1615;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 1615 BP; 321 A; 458 C; 493 G; 343 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 210.4; DB 22;
Pred. No. 1.2e-49;
0; Mismatches 256;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                at ftp.wipo.int/pub/published_pct_sequences
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; SED ID No 1074; 546pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Barash SC, Ruben SM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (HUMA-) HUMAN GENOME SCI INC.
                                                                                                                                                                                                                                                                                                                                                                                                                  2000US-0251989.
2000US-0251990.
2000US-0254097.
                                                                                                                                                                                                                                                                                                                 2000US-0251030.
2000US-0251988.
                                                       2000US-0249210
                                                                       2000US-0249211
                                                                                 2000US-0249212
                                                                                                  2000US-0249213
                                                                                                                              2000US-0249215
                                                                                                                                          2000US-0249216
                                                                                                                                                                                                                                                                          2000US-0249300.
                                                                                                                                                                                                                                                                                                    2000US-0250391.
                                                                                                                                                                                                                                                                                                                                            2000US-0256719
                                                                                                                                                                                                                                                                                                                                                           2000US-0251479
                                                                                                                                                                                                                                                                                                                                                                                                      2000US-0251869.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           2001US-0259678.
                            2000US-0249208
                                                                                                               2000US-0249214
                                                                                                                                                         2000US-0249217
                                                                                                                                                                                    2000US-0249244
                                                                                                                                                                                                    2000US-0249245
                                                                                                                                                                                                                  2000US-0249264
                                                                                                                                                                                                                                                           2000US-0249299
                                                                                                                                                                                                                                                                                      2000US-0250160
                                                                                                                                                                                                                                            2000US-0249297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2001-476224/51.
                                      17 - NOV - 2000;
                                                                                                                                                                                                               17 - NOV - 2000;
17 - NOV - 2000;
17 - NOV - 2000;
17 - NOV - 2000;
17 - NOV - 2000;
10 - DEC - 2000;
05 - DEC - 2000;
06 - DEC - 2000;
06 - DEC - 2000;
08 - DEC - 2000;
08 - DEC - 2000;
                                                                                                            17-NOV-2000;
17-NOV-2000;
17-NOV-2000;
17-NOV-2000;
17-NOV-2000;
17-NOV-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           05-JAN-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                               08-DEC-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Rosen CA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
```

```
Human colon cancer cell line polynucleotide sequence SEQ ID NO:2361
                                                                     Human; colon cancer; tumour; diagnosis; gene expression product; probe; detection; cancerous state; metastesis; identification; breast cancer; oestrogen receptor-positive breast cancer; therapy; oestrogen receptor-negative breast cancer; lung cancer; ss.
                                                                                                                                                                                                       98US-0085426.
98US-0085537.
98US-0085696.
98US-0105234.
                                                                                                                                                                                                                                                                     (CHIR ) CHIRON CORP (HYSE-) HYSEQ INC.
                                                                                                                                                                                                                  15-MAY-1998;
15-MAY-1998;
21-OCT-1998;
27-OCT-1998;
```

1;

Gaps

3;

Indels

1384 gggctgtaggctttgtgcatagcccacttctgaaaaacaagggaacagtgtgtaaggaac 1443

ö

Conservative

Local Similarity

Matches 377;

à g

2000US-191637P. 2000US-0614150.

23-MAR-2000; 11-JUL-2000;

23-MAR-2001; 2001WO-US09231

us-09-495-823-8.rng

```
4;
                                                                                                                                                                                                                                                         genes correlated with a cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample derived from a cell suspected of being cancerous, where detection of the differentially expressed gene product is correlated with a cancerous state of the cell from which the test sample was derived. The polynucleotides sequences can be used in a method for detecting differentially expressed genes correlated with a cancerous state of mammalian cell. The polynucleotides can also be used as probes for detecting and mapping related genes. They can be used in diagnosis and prognosis of diseases and disorders (e.g. identification of
                                                                                                                                                                                                          AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
libraries constructed from human colon cancer cell lines. The present
invention also describes a method of detecting differentially expressed
                                                  ï
                                                                                                                              Polynucleotide library used to determine cancerous states of mammalian
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gtettegeteacceegagtagatattttgeataaca-ttgaceceatatacaccaaggea 1610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      aaaaatggctcctgggcagcagctat-gggatctggaacactgcaatccagtcagccat 1669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   acagattgatgaggacattcaacta-gatggctatgatatctgggagaccataagtgagg 1551
                                                                                                                                                                                                                                                                                                                                                                                                                               prefmetastatic or metastatic cancerous states, stages of cancer, or responsiveness of cancer to therapy). This is particularly for breast cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    acagattgatgaggacattcaactaggatggctatgatatctgggagaccataagtgang 327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                 Sudduth-Klinger J;
                                  Kassam A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Drosophila melanogaster expressed polynucleotide SEQ ID NO 33059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ,
,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 21; Length 1082;
Escobedo J, Innis MA, Garcia PD, Junessai
Glese K, Randazzo F, Kennedy GC, Pot D, Kassai
manac R, Crkvenjakov R, Dickson M, Drmanac S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 1082 BP; 230 A; 182 C; 155 G; 190 T; 325 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   negative breast cancer, lung cancer, and colon cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 139.6; DB 21;
Pred. No. 2.1e-29;
0; Mismatches 9;
                                                                                                                                                                             Claim 1; Page 938; 1097pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1670 -- cagagtgcagcactggaaa 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ABL12859 standard; cDNA; 1740
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 93.0°
Matches 187; Conservative
          d C, Giese K, G, Drmanac R, F, ta D,
                                                              Leshkowitz D, Kita D,
                                                                                             WPI; 2000-126369/11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     26-MAR-2002
               Williams LT,
Reinhard C,
                                                Lamson G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABL12859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABL12859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    음
```

developmental biology; cell signalling; insecticide;

Drosophila melanogaster.

WO200171042-A2

27-SEP-2001

pharmaceutical; gene;

Drosophila;

```
The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                    detection reagent for detecting 1000 or more for elucidating cell signalling and cell-cell
                                                                                                                                                                                                                                                                                                                                                capable of detecting 1000 or more genes from Drosophila. The invention useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gtcctgggatgtgtggctatgacttgtatgaaaacgacaatgctg---cctgggactatg 1020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            acaatggcatatactccacacagatgtacactcagagagtacagcaaatcttagcttccc 1080
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      140 gotogaatoagatoottacaccgaatatagatgotttggcotacaatggtatootgttga 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           723
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    783
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            843
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           963
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 560 cgaatggcacttacgccaccgaagcctttacttcggaagcaaaaaggataattgagcaac 619
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              aaccaaatatagttataatactaatcgatgacatgggcatgaacgatgtcagttttcatg 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                 The invention relates to an isolated nucleic acid detection reagent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           agccccatctcattttcatcctagcggatgatcagggatttagagatgtgggttaccacg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gatc---tgagattaaaacacctactcttgacaagctcgctgccgaaggagttaaactgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ataagcattatgttccgaacctatgcacaccatccagagctactctgctcactggaaaat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    atcagatacacaccoggacttcaacattctatcataaqacctacccaacccaactgtttac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ateceatacacagagatgeaggactttgtgattateacegatgagecgtggggeette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ctcagcgagaacgtcttatgcccgaaattttccgggatgcgggctactcgacccatttgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               atacctttttttggttcccttttgggaagtggggattactatacacactacaaatgtgaca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ctctggacaatgccaccctacctcagaaactgaaggaggttggatattcaacgcatatgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   tgggcaagtggcacttgggcttctggcgcaaggatctcacgccaacaatgcgtggattcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              agaactactatgtccagcctatttgcacaccatccaggagtcagtttattactggaaagt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         tegataggaactacteggetggactggattteeggeggggatetggaggeegtgeeeggaag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1740;
                                                                                                                                                                                                                                                                                            ID NO 33059; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 1740 BP; 469 A; 420 C; 441 G; 410 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 129; DB 23; Pred. No. 3.1e-26; O; Mismatches 500;
                                                                                                                             Myers EW;
                                                                                                                             PWD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 5.7%;
Best Local Similarity 48.5%;
Matches 502; Conservative (
                                                                                                                                                                                                                      New isolated nucleic acid
genes from Drosophila and
                                                                                                                             Ľ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (ABB57737-ABB72072).
                                                                                                                             Venter JC, Adams M,
                                                                                                                                                              WPI; 2001-656860/75.
P-PSDB; ABB68756.
                                                                                         (PEKE ) PE CORP NY
                                                                                                                                                                                                                                                           interactions
                                                                                                                                                                                                                                                                                                Claim 1; SEO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           904
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                440
δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         pp
```

4

```
1081 ataaccccacaaagcctatatttttatatattgcctatcaagctgttcattc----- 1132
                      1133 -----accactgcaagctcctggcaggtatttcgaacactaccgatccattatcaaca 1185
                                                                                                                                               980 ttaaggagegaggatatgtgtccaaccaggegatacatgccgtagactggetgccacac 1039
       680 aggatagtccgatgcaggctccagaggaagtggccaaatttccccacatcagggatc 739
                                                                                                                                                                                               gggagaccataagtg 1548
                                                                                                                                                                                                       || | | ||||||
ggcctatgctgagtg 1114
                                                                                                                                                                                               1534
                                                                                                                                               1414
                                                                                            ò
                                                                                                         qq
                                                                                                                   g ç
                                                                                                                                               ò
                                                                                                                                                          Q
                                                                                                                                                                     ò
                                                                                                                                                                                 q
                                                                                                                                                                                              à
à
          g
                       ö
                                  a
                                            ŏ
                                                         q
                                                                    ò
                                                                                a
```

Search completed: July 23, 2002, 17:07:11 Job time: 7302 sec

```
GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 23, 2002, 17:15:04 ; Search time 69.6 Seconds

(without alignments)

759.327 Million cell updates/sec
```

Title: US-09-495-823-7
Perfect score: 3012
Sequence: 1 MAPRGCAGHPPPPSPQACVC......VPVRYPPKDPRSNPRLNGGV 550
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters:

283138 segs, 96089334 residues

Searched:

283138

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : PIR\_71:\*

e: PIR\_71:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

N-acetylgalactosam N-acetylgalactosam Sulfatase family p probable secreted arylsulfatase PA01 N-acetylgalactosam arylsulfatase (EC steryl-sulfatase (EC arylsulfatase E (E arylsulfatase E (E arylsulfatase E (E arylsulfatase E (E arylsulfatase [imp arylsulfatase [imp arylsulfatase [imp arylsulfatase (EC probable arylsulfa arylsulfatase (EC probable arylsulfatase (EC probable arylsulfatase (EC probable arylsulfatase (EC probable arylsulfatase (EC hypothetical prote probable arylsulfa probable sulfatase arylsulfatase homo probable sulfatase cerebroside-sulfat arylsulfatase (EC arylsulfatase (EC probable sulfatase N-acetylgalactosam arylsulfatase (EC Description SUMMARIES A44475 154210 AF07394 AF07506 DB3622 DB3622 KJHUG6 KJHUG6 I37362 S69336 I37187 I37187 I37187 E86066 E86066 D95377 T45548 B35159 AF0103 T15835 F95325 KJHUAB S07089 S30691 DB Length Query 1424.5 1414.5 1724.4 472.4 472.4 472.4 472.5 472.5 407.5 398.3 3 398.3 3 398.3 398.3 398.3 398.3 398.3 398.3 398.3 398.3 Score Result NO. 

probable sulfatase probable sulfatase	choline-sulfatase	probable sulfatase	hypothetical prote	N-acetylglucosamin	arylsulfatase (EC	probable sulfatase	probable sulfatase	probable sulfatase	probable sulfatase	choline sulfatase	hypothetical prote	probable sulfatase	probable sulfatase	iduronate-2-sulfat
B70643 AF0370	AD3523	AB0102	T29618	KJHUGU	S43229	E70533	C91206	E86052	G65169	E83642	T16584	B70535	F83354	KJHUID
7 7	7	~	7	٦	7	7	~	~	N	~	7	7	7	Т
787 557	495	517	452	552	649	970	497	497	497	503	109	787	538	550
10.6 10.4	0.6	8.5	8.5	8.3	8.2	8.1	8.1	8.1	8.0	7.4	7.3	7.1	7.0	6.8
319.5	272	257.5	256	248.5	247	243.5	243	243	241	221.5	219.5	213.5	212	205.5

## ALIGNMENTS

3 3 3 5 6 7 7 6 7 7 6 7 7 6 7 7 6 7 8 9 7 8 9 7 7 7 6 9 1 9 1 9 1 9 1 9 1 9 1 9 1 9 1 9 1 9	KJHUAB	
13. Dec-1993 #teaxt_change 08-Dec-2000 13. Dec-1993 #sequence_revision 27-Oct-1995 #teaxt_change 08-Dec-2000 100: \$33907; \$3307; \$435078; \$4449; \$84449; \$642449; \$154217; \$45689; \$40.0 M.B. 14 to the EMBL Data Library, March 1993 100: \$35990 100: \$35990 100: \$35990 100: \$35990 100: \$35990 100: \$35990 100: \$35990 100: \$35990 100: \$35990 100: \$35990 100: \$35990 100: \$35990 100: \$35990 100: \$35990 100: \$35990 100: \$4577 100: \$4577 100: \$45090 100: \$4577 100: \$45090 100: \$4577 100: \$45090 100: \$45090 100: \$4577 100: \$45090 100: \$45000 100: \$450	N-acetylgalactosamine-4 sulfatase (EC 3.1.6.12) N.Alternate names: arylsulfatase B (ASB); chondr	.;
d to the EMBL Data Library, March 1993  note number: 835990  an 1-533 <-PET    set 1-533 <-PET    terferences: EMBL:X72735, NID:9289009, PIDN:CA551272.1; PID:9825628  set 1-533 <-PET    set 1-534 <-PET    set 1-534 <-PET    set 1-535 <-PET    set 1-535 <-PET    set 1-536 <-PET	C; Specific nomic Sapiens (man) C; Date: 31-Dec-1993 #sequence_revision 27-Oct-19 C.Accession. S75090. S33077. A35078. A45550. A42	
a to the EMBL Data Library, March 1993  noe number: 835990  ton: 835990  ton: 835990  ton: 835990  ton: 835900  ton: 835000  ton: 83500	R; Peters, C.W.B.	
ion: S3590 le type: DNA se: 1.533 cPET> se: 1.533 cPET> seferences: EMBL:X72735; NID:g289009; PIDN:CAA51272.1; PID:g825628 ssi, S.: Rupp, K.: Von Figura, K.; Peters, C. am. Hoppe-Seyler 374, 327-335, 1993 am. Hoppe-Seyler 374, 327-335, 1993 am. Hoppe-Seyler 374, 327-335, 1993 am. Hoppe-Seyler 374, 327-335, 1993 cloning and expression cloning and expressio	submitted to the EMBL Data Library, March 1993 A;Reference number: S35990	
Target: Manage Septembers: Rubb. X72735; NID: 9289009; PIDN: CAA51272.1; PID: 9825628  Set S.; Rupp, K.; von Figura, K.; Peters, C.  Set S.; Rupp, K.; von Figura, R.;  Set Septembers: And Se	A; Accession: S35990	
references: EMBL:X72735; NID:9289009; PIDN:CAA51272.1; PID:9825628 Ssi, S.; Ruppe, K.; Von Figura, K.; Peters, C.  A. Hoppe-Seyler 374, 327-335, 1993  A. Hoppe-Seyler 374, 327-336, 1993  A. Hoppe-Seyler 374, 327-336, 1993  ASSI  PASSI  RABL:X72738; EMBL:X7273  Isdorf, M.; Vingron, M.  Cloning and expression  Cloning and expression  PID:9179077  Hopwood, J.J.  maturation and isolatic  Mucopolysaccharidosis type VI: identification or rehere metations in the a noce number: A42449; MUID:92197625  ion: A42449; MUID:92197625  ion: A42449; MUID:92197625	A;Molecule type: DNA A:Residues: 1-533 <pft></pft>	
Ssi, Si, Rupp, K.; Von Figura, K.; Peters, C.  man. Hoppe-Seyler 374, 327-335, 1993  man. Hoppe-Seyler 374, 327-335, 1993  lsdorf, M.; Vingron, M.  lsdorf, M.; Vingron, M.  cloning and expression  i. PID:g179077  equencing  Hopwood, J.J.  maturation and isolatic  maturation and isolatic  mature protein, were  mucopolysaccharidosis type VI: identification or three mutations in the a  nce number: A42449; MUID:92197625  lion: A42449; MUID:92197625	A; Cross-references: EMBL:X72735; NID:g289009; PI	N:CAA51272.1; PID:9825628
(02 C5)	R;Modaressi, S.; Rupp, K.; von Figura, K.; Peter Biol. Chem. Hoppe-Sevler 374, 327-335, 1993	O.
102 (5)   >45%   EMBL:X72738; EMBL:X72738; EMBL:X72733   Isdorf, M.; Vingron, M. cloning and expression cloning and expression   : PID:g179077   equencing   Hopwood, J.J.     Hopwood, J.J.	A; Title: Attention of the himan arelenifatase B	iene.
EMBL:X72738; EMBL:X72737  Isdorf, M.; Vingron, M.  cloning and expression  cloning and expression  i. PID:g179077  equencing  Hopwood, J.J.  maturation and isolatic  i. PID:g236688  i. PID:g236888  i. PID:g236888  i. PID:g236888  i. PID:g	A; Refere	
EMBL:X72738; EMBR:X72738; EMBL:X72738; EMBL:	A; Access	
EMBL:X72738; EMBL:X72737  Isdorf, M.; Vingron, M.  cloning and expression  ; PID:g179077  equencing  Hopwood, J.J.  maturation and isolatic  NCBID:g236698  NCBID:577789  the mature protein, were necessidue  whoropolysaccharidosis type VI: idemtification-or-three mutations in the ance number: A42449; MUID:92197625  ie type: mRNA	(62)	
lsdorf, M.; Vingron, M.  cloning and expression ; PID:g179077 equencing Hopwood, J.J. maturation and isolatic maturation and isolatic ,; PID:g236698 NCBID:57778) the mature protein, were Mucopolysaccharidosis type VI: idemtification or residue nce number: A42449; MUID:92197625 he type: mRNA	10 10	EMBL: X72738; EMBL: X72739;
Isdorf, M.; Vingron, M.  cloning and expression ; PID:g179077 equencing Hopwood, J.J. maturation and isolatic modelysaccharidosis type VI: identification-or-three mutations in the 6 noc number: A42449; MUID:92197625 le type: mRNA	A; Note:	
: PID:g179077 equencing Hopwood, J.J. maturation and isolatic maturation and isolatic mopolysaccharidosis type VI: identification-or-three mutations in the a nce number: A42449; MUID:92197625 le type: mRNA	R; Peter:	
; PID:g179077 equencing Hopwood, J.J. maturation and isolatic modolysaccharidosis type VI: identification or the mature protein, were number: A42449; MUID:92197625 le type: mRNA	J. B101	or o
; PID:g179077 equencing Hopwood, J.J. maturation and isolatic "YCBIP:577789  The mature protein, were whore between residue noe number: A42449; MUID:92197625 le tvoe: mRNA	A; Befer	and expression
; PID:g179077 equencing Hopwood, J.J. maturation and isolatic maturation and isolatic "VEBIP:57778) the mature protein, were newhere between residue noc number: A42449; MUID:92197625 le tvoe: mRNA	A; Acces:	
; PID:g179077 equencing Hopwood, J.J. maturation and isolatic .; PID:g236698 NCBID:57778) the mature protein, were lewhere between residue nce number: A42449; MUD:92197625 le tvoe: mRNA	A; Moleci	
; PID:91/90// equencing Hopwood, J.J. maturation and isolatic .; PID:9236698 NCBIP:5778) the mature protein, were mucopolysaccharidosis type VI: identification-or-three mutations in the 6 non Malay MUID:92197625 le type: mRNA	A; Resid	
equencing  Hopwood, J.J.  maturation and isolatic  "SELP:57778)  the mature protein, were much protein, were now protein, which were now protein, which were now protein, were now protein, which were now protein, were now protein, which were not not not not not not now protein, which were not	A; Cross	; PID:q179077
Hopwood, J.J.  maturation and isolatic  ", PID:9236698  "CBIP:57778)  the mature protein, were  "Mucopolysaccharidosis type VI: identification or three mutations in the a nce number: A42449; MUID:92197625 le type: mRNA	A; Note:	equencing
Hopwood, J.J.  maturation and isolatic  .; PID:g236698 NCBIE:57778)  the mature protein, were mucopolysaccharidosis type VI: identification or "three mutations in the annumber: A42449; MUID:92197625 le type: mRNA	A; Note:	
maturation and isolatic  "" PID:g236698  NCBIP:57778)  the mature protein, were mucopolysaccharidosis type VI: identification-or-three mutations in the and a type of the mature mature of the and the angle of the a	K;Litje	Hopwood, J.J.
", PID:9236688  "NCBIP:57778)  the mature protein, were much mature protein, were now mature mature mature in the 6 mature	D.T.i+10	
.; PID:9236698 NCBIP:57778) the mature protein, wer newhere between residue nce number: A42449; MUID:92197625 le type: mRNA	A:Refer	
.; PID:9236698 NCBIP:57778) the mature protein, wer where between residue since number: A42449; MUD:92197625	A; Acces.	
.; PID:9236698 NCBIP:57778)  the mature protein, were mucopolysaccharidosis type VI: identification or three mutations in the and the mature mutations in the answer to the mature mature in the answer management of the mature m	A; Molec	
.; PID:9236698 NCBIP:57778) the mature protein, were mucopolysaccharidosis type VI: identification or three mutations in the series number: A42449; MUD:92197625 lie type: mRNA	A; Resid	
NCBIP:57778)  the mature protein, wer iewhere between residue : Mucopolysaccharidosis type VI: identification or three mutations in the & sion: A42449, MUID:92197625	A; Cross	.; PID:g236698
the mature protein, were mucopolysaccharidosis type VI: identification or three mutations in the since number: A42449; MUD:92197625	A; Note:	NCBIP: 57778)
the mature protein, wer immorphysaccharidosis type VI: identification or three mutations in the ance number: A42449; MUID:92197625	A; Note:	
lewhere between residue: Mucopolysaccharidosis type VI: identification or three mutations in the sion: A12449; MUID:92197625 sion: A12449	A; Note:	the mature protein, were d
R:JIN, Am. J	A; Note:	newhere between residue 450
Am. J	R;Jin,	
eity. A; Reference number: A42449; MUID:92197625 A; Accession: A42449 A:MOID:0019 type mRNA	Am. J. A. Title: Mucopolysaccharidosis type VI: identifi	Sation or three mutations in the aryl
A; Reference number: A42449; MUID:92197625 A; Accession: A42449 A:Molecule tone: mRNA	eity.	•
A.M.C.C.S.J.LOL. A.M.Z.4.4.7 A.M.O.J.P.C.I.J.P. T.V.D.P. B.R.N.A.	A; Reference number: A42449; MUID:92197625	
	A.Molecule type: mRNA	

```
Natering alactosamine-4-sulfatase (EC 3.1.6.12) precursor - cat
N.Alternate names: arylsulfatase B (ARSB); chondroitinase; chondroitinsulfatase; G. Specials: Fellia silvestris catus (domestic cat)
C. Specials: Fellia silvestris catus (domestic cat)
C. Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C. Accession: A44475
R. Jackson, C.E.; tuhki, N.; Desnick, R.J.; Haskins, M.E.; O'Brien, S.J.; Schuchman, E Genomics 14, 403-411, 1992
A; Title: Feline arylsulfatase B (ARSB): isolation and expression of the cDNA, compari A; Reference number: A44475; MUID: 93052342
A; Accession: A44475
A; Status: preliminary; not compared with conceptual translation
A; Accession: A44475
A; Status: preliminary; not compared with conceptual translation
A; Molecule type: mRNA
A; Residues: 1-535 < JAC>
A; Cross-references: GB: S48472; NID: 9258855; PIDN: AAB23941.1; PID: 9258856
A; Cross-references: GB: S48472; NID: 9258855; PIDN: AAB23941.1; PID: 9258856
A; Cross-references: GB: S48472; NID: 9258855; PIDN: AAB23941.1; PID: 9258856
A; Cross-reference extracted from NCBI backbone (NCBIP: 117976)
C; Comment: This enzyme is frequently misidentified as EC 3.1.6.1.
A; Description: hydrolyzes N-acctylgalactosamine-4-sulfate units in chondroitin sulfatase
C; Superfamily: animal sulfatase
C; Superfamily: animal sulfatase
C; Reywords: glycoprotein lysosomal storage disease; lysosome; sulfuric ester hydrola F; 1-42/Domain: signal sequence #status predicted <GMATP
F; 1-42/Domain: signal sequence #status predicted <GMATP
F; 466-535/Product: apma chain #status predicted <GMATP
F; 190, 281, 428, 460/Binding site: carbohydrate (Asn) (covalent) #status predicted
F; 190, 281, 428, 460/Binding site: carbohydrate (Asn) (covalent)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6
209 MYSTNIFTKRAIALITNHPPEKPLFLYLALQSVHEPLQVPEEYLKPYDFIQDKNRHHYAG 268
                                                                                                                                                                                                                                                                                                                                                                             415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            389 IELLHNIDPNFVDSSPCPRNSMAPAKDDSSLPEYSAFNTSVHAAIRHGNWKLLTGYPGCG 448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          465 DWVPPQSFSNLGPNRWHNERITSS--TGKSVWLFNITADPYERVDLSNRYPGIVKKLLRR 522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 416 VDILHNIDPIY-----TKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGNPGYS 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EGALLAQAGEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLE 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    54
                                                                                                                                                                                                                                                                                                                                                   MLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 MAPRGCAGHPPPPSPQACVCPG-----KMLAMGALAGFWILCLLTYGYLSWGQALEEEE
                                                                                                                                                                                                                       236 IYSTQMYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRSIININRRRYAA
                                                                                              FYRKECMPTRRGFDTFFGSLLGSGDYYTHYKCD----SPGMCGYDLYENDNAAWDYDNG
                                                                                                                                                         150 MYRKECLPTRRGFDTYFGYLLGSEDYYSHERCTLIDALNVTRCALDFRDGEEVATGYKN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VGFVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     47.0%; Score 1414.5; DB 1
49.0%; Pred. No. 1.8e-100;
iive 82; Mismatches 143;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      523 LSQFNKTAVPVRYPPKDPRSNPRLNG 548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | ::| :|| :|| :| :|| :|| :|| || LQFYHKHSVPVYFPAQDPRCDPRATG 527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches 280; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local
                                                                                                                                                                                                                                                                                                                                                   296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            356
                                90
                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                              569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 502
                                                                                              δ
                                                                                                                                                      qq
                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                           A; Wolecule type: mRNA
A; Residues: 403-404, TY', 406-407 <J13>
A; Residues: 403-404, TY', 406-407 <J13>
A; Residues: 403-404, TY', 406-407 <J13>
A; Residues: 403-404, TY', 406-407 <J13-40;
A; Cross-references: GB:S90743; NID:9247490; PIDN:AAB21833.1; PID:9247491
A; Note: the enzyme is referred to as EC 3.1.6.1
A; Note: these mutations give rise to mucopolysaccharidosis type VI, Maroteaux-Lamy disea
B; Schmidt, B; Selmer, T; Ingendoh, A; von Figura, K.
Cell 82, 271-278, 1995
A; Title: A novel amino acid modification in sulfatases that is defective in multiple sul
A; Reference number: A57113; MUID:95394208
A; Contents: annotation; identification of 3-oxoalanine, 2-amino-3-oxopropanoic acid
B; Schuchman, E.H.; Jackson, C.E.; Desnick, R.J.
Genomics 6, 149-158, 1990
A; Title: Human arylsulfatase B: MOPAC_cloning, nucleotide sequence of a full-length CDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     human place
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          dis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οį
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      beta chains
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ٠.
ش
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-357, 'W', 377-533 <RES-
A;Residues: 1-357, 'W', 379-375, 'M', 377-533 <RES-
A;Cross-references: GB:M32373; ND:9179029; PIDN:AAA51779.1; PID:9179030
R;Kobayashi, T.; Honke, K.; Jin, T.; Gasa, S.; Miyazaki, T.; Makita, A.
Biochim. Biophys. Acta 1159, 243-247, 1992
A;Title: Components and proteolytic processing sites of arylsulfatase B from 1A;Reference number: A56865; MUID:93003385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QAGEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQP 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121 ICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLA 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             34 PPGS----GAGASRPPHLVFLLADDLGWNDVGFHGSRIRTPHLDALAAGGVLLDNYYTQP 89
                                                                                                                                  A; Residues: 234-235, 'P', 237-238 <112>
A; Cross-references: GB:S90736; NID:g247488; PIDN:AAB21832.1; PID:g247489
A; Note: sequence extracted from NCBI backbone (NCBIN:90736, NCBIP:90739)
A; Accession: C42449
                             Cross-references: GB:S90729; NID:9247486; PIDN:AAB21831.1; PID:9247487 Note: sequence extracted from NCBI backbone (NCBIN:90729, NCBIP:90731)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Map position: 5q11-5q13
A;Introns: 104/3; 167/1; 230/3; 300/1; 381/2; 405/1; 446/1
A;Note: defects in this gene can cause mucopolysaccharidosis type VI,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A Molecule type: protein
A; Residues: 41-55,444,425, X',427-454;466-483 <KOB>
A; Experimental source: placenta
A; Note: sequence modified after extraction from NCBI backbone
A; Note: the fragments shown are the amino ends of the alpha, gamma,
C; Comment: This enzyme is frequently misidentified as EC 3.1.6.1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                47.3%; Score 1424.5; DB 1; Length 533;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. 3e-101;
; Mismatches 149;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Cross-references: GDB:119008; OMIM:253200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Reference number: 154217; MUID:90152677
A; Accession: 154217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    81;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          49.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 49.39
Matches 279; Conservative
                                                                                           A; Accession: B42449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Gene: GDB: ARSB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Function:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

39AGAGADRPPHLVFVLADDLGWNDVSFHGSNIRTPHLDELAAGGVLLD 115 NYYVOPICTPSRSQFITGKYOIHTGLOHSIIRPTQPNCLPLDNATLPOKLKEVGYSTHMV	Qy         269         HSPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDN 328           Db         182         HDPLQVPEEYMEPYDFIQDKHRRIYAGMYSLLDEAVGNVTRALKSRGLWNNTVLIFSTDN 241           Qy         329         GGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNGTYCEFVHTTDWYPTLISLA           Qy         329         GGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNGTYCEFWHTTDWYPTLISLA           Qy         389         EGQDEDIQLDGYDIWETISEGLRSPRYDILHNIDPTY           Qy         389         EGQTBEDIQLDGYDIWETISEGLRSPRYDILHNIDPTY           A38         FGGTRSGGNWPLTGREGSPRYDILHNIDPTY           Qy         438         YGIWNTAIOSAIRVQHWKLLTGNPQYSDWVPPQSFSNLG-PNRWHNERITSSTGKSVWL           Qy         438         YGIWNTAIOSAIRVQHWKLLTGNPQYSDWVPPQSFSNLG-PNRWHNERITSSTGKSVWL           Qy         438         YGIWNTAIOSAIRVWKLLTGYPGCGYWFPPPSQSNISEVPSVDSPTKTLWL           Qy         436         FNITADPYPRYBUZNRYKKLRRELGSPNYRTAVPPYRPRENPSNPRLMG         548           QS         HSAFNTSIHAGIRNYRTHRIYQULLSRLQYYHENSVPSYPPLDPRCDPKGTG         467
OY 410 GLRSPRUDILHNIDDIYTRAKNGSWAAGYGIWNTAIQSAIRVQHWKLLT 458  OY 410 GLRSPRUDILHNIDDIYTRAKNGSWAAGYGIWNTAIQSAIRVQHWKLLT 458  1   1	RESULT 4  H87394  Sulfatase family protein [imported] - Caulobacter crescentus C;Species: Caulobacter crescentus C;Species: Caulobacter crescentus C;Species: Caulobacter crescentus C;Accession: H87394 R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B; Laub, M.T.; DeBoy, K.T.; DodSon, K.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001 A;Title: Complete Genome Sequence of Caulobacter crescentus. A;Reference number: A87249; WUID:21173698; PMID:11259647
RESULT 3 154210 N-acetylgalactosamine-4-sulfatase (EC 3.1.6.12) precursor - rat (fragment) N.Alternate names: arylsulfatase (EC 3.1.6.1) B [misidentification] C;Species: Rattus norvegicus (Norway rat) C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 02-Jun-2000 C;Accession: 154210	A;Accession: H87394 A;Status: preliminary' A;Solatus: preliminary' A;Molecule type: DNA A;Residues: 1-521 <sto> A;Cross-references: GB:AE005673; NID:g13422494; PIDN:AAK23156.1; GSPDB:GN00148 C;Genetics: A;Gene: CC1172</sto>
R.Kunieda, T. R.Kunieda, T. R.Kunieda, T. S. S12-587, 1995 A.Title: Mucopolysaccharidosis type VI in rats: Isolation of CDNAs encoding arylsulfatas A; Reference number: 154210 A; Recession: 154210 A; Status: preliminary; translated from GB/EMBL/DDBJ A; Molecule type: mRNA A; Status: 1-473 < RES> A; Cross-references: GB: D49434; NID: 91065603; PIDN: BAA08412.1; PID: 91089794 C; Genetics: A; Gene: ARSB	Duery Match 15.7%; Score 472; DB 2; Length 521; Best Local Similarity 26.6%; Pred. No. 3.2e-28; Matches 142; Conservative 69; Mismatches 193; Indels 130; Ga 68 PSTTSTSQPHLIFILADDGFRDVGYHGSEIKTPTLDKLAAEGVKLEN-YYVQPI
;	Db 85 CAPSRAAIMTGRYATREGEEFTPTPVAESRVVGGHAGDPLHPSRENGAEVKNMEKDENVL 144  Qy 152 CLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDTFFGSLLGSGD 205
AACGVKLENYYQPICTPSRSQFITG AACGVVLDNYQPLCTPSRSQLITG AGGVVLDNYQPLCTPSRSQLLITG VGYSTHWYGKWHLGFYRKECMPTRR 	Db 204

QO	368 EQLPAGAVYRSPVGHVDIFATAAGAAGAPTPKDRVLDGVDLVPFVKGQSSG 418	
Qy Dp	423 DPIYTKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGNPGYSDWVPPQSFSNLGPNRWHN 482 	RESULT 6 D83622 arylsulfatase PA0183 [imported] - Pseudomonas aeruginosa (strain PA01)
Oy Dp	483 ERITSSTGKSVWLFNITADPYERVDLSNRYPGIVKKLLRRLSQFNKTAVPVRYP 536     :::             :::       450TWLPDLSTDPTERQELSKARPEKLREMQALMAQLDGQMMKPTWP 493	<pre>C.Species: Pseudomonas aeruginosa C.Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000 C.Accession: D83622 R.Stover, C. K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L.</pre>
RESULT AF0506 probabl C; Spec A; Note	06 able secreted sulfatase STY0042 [imported] - Salmonella enterica subsp. enterica ser ecies: Salmonella enterica subsp. enterica serovar Typhi te: this species has also been called Salmonella typhi te: this species has also been called Salmonella typhi te: the species has also been called Salmonella typhi te: the species has also been called Salmonella typhi	quence of Pseudomonas aeruginosa PAO1, an opportunist MUID:20437337
C; A( C; A( th, th, S.	C. Accession: AP6506 R. Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.	A. Residues: 1-536 CSTO> A. Residues: 1-536 CSTO> A. Cross-references: GB. AE004456; GB. AE004091; NID: 99946013; PIDN: AAG03573.1; GSPDB:GN A. Experimental source: strain PA01 C. Genetics:
Nati A; Ai A; Ti	re 413, 848-852, 2001 Lhors: Parry, C.; Quall, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Lle: Complete genome sequence of a multiple drug resistant Salmonella enterica serov ference number: ABO502: PMID:11677608	A,Gene: atsA; PAO183 C;Superfamily: animal sulfatase
A; Ac A; St A; Mc	A;Accession: AF0506 A;Status: preliminary A;Molecule type: DNA a. postding: 1.407 cnbb.	Query Match 14.2%; Score 429; DB 2; Length 536; Best Local Similarity 24.9%; Pred. No. 6.5e-25; Matches 144; Conservative 71; Mismatches 190; Indels 174; Gaps 20;
\$ C \$ S	Jacues: 1-45/ New John Capoll90.1; PID:g16501319; GSPDB:GN00176 netics: STY0042	Qy 73 TSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAABGVKLENYYVQPICTPSRSQFITG 132 : : :: ::
O B N	Query Match 15.1%; Score 454; DB 2; Length 497; Best Local Similarity 26.3%; Pred; No. 7.18-27; Astronomy 170, Astronomy 190, Astronomy 170, Astronomy 190,	QY 133 KYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKW 177
7 6	PHILIFILADDQGFRDVGYHGSEI-KTPTLDKLAABGVKLENYYV- 118  DANIVITININ	QY 178 HLGFYRKECMPTRRGFDTFFGSLLGSGDYTHYKCDSPGMCGYDLYEN 225     :
3 6	119 QPICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMYGKWH 178   119 QPICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMYGKWH 178   1	OY 226 DNAAWD-YDNGIYSTOMYTORVQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRS 284 
3 6	179 LGEYRKECMPTRRGFDTFFGSLLGSGDYTHYKCDSPGMCGYDLYENDNA 228  135 LGEYRKECMPTRRGFDTFFGSLLGSGDYTHYKCDSPGMCGYDLYENDNA 228  135 INAGCEPPHOLOGOAN CHARLED CONTRACTOR C	QY 285RRYAAM 296 :::  HIII Db 227 RYDAGPEALRQERLARLKELGLVEADVEAHPVLALTREWEALEDEERAKSARAMEVYAAM 286
3 6 8	IFLYIAYQAVHSPLQAPG 27	QY 297 LSCLDEAINNVTLALKTYGFYNNSIIIXSSDNGGQPTAG 335 :
3 6	277 RYFEHYRSIININRRYAAMLSCLDEAINNYTLALKTYGFY 317 278 RYFEHYRSIININRRYAAMLSCLDEAINNYTLALKTYGFY 317 238 RYFDAVSCOVARDVORGIDII EVERGAARDERINDERINDERINDERINDERINDERINDERINDERI	QY 336GSNWPLRGSKGTYMEGGIRAVGFVHSPLLKNKGTVCKEPVH 376  Db 347 LENIGRANSYVWYGPRWAQAATAPSRLYKAFTTQGGIRVPALVKYPRLSRQGAISHAFAT 406
3 6	NNSIIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGSVG	QY 377 ITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPIYTKAKNGS 433 :
3 3 8		QY 434 WAAGYGIWNTAIQSAIRVQHWKLLTGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGKSV 493  458 -TGWELFGMRAIRQCDWKAVYLPAPVGPATWQ
Qy Dp	IYTKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGNPGYSDWVPPOSFSNLGPNRWHNER 48  PRQDDPTDEWAIRDGDWKM43	Oy 494 WLFNITADPYERVDLSNRYPGIVKKLLRRLSQF-NKTAV 531 
Qy Dp	485 ITSSTGKSVWLFNITADPYERVDLSNRYPGIVKKLLRRLSQFNK 528 	RESULT 7 KJHUG6 N-acetylgalactosamine-6-sulfatase (EC 3.1.6.4) precursor [validated] - human

```
Arylsulfatase (EC 3.1.6.1) precursor - sea urchin (Strongylocentrotus purpuratus)
C; Species: Strongylocentrotus purpuratus (purple urchin)
C; Species: Strongylocentrotus purpuratus (purple urchin)
C; Species: Strongylocentrotus purpuratus (purple urchin)
C; Saccession: A37362
R; Yang, O.; Angerer, L.M.; Angerer, R.C.
Dev. Biol. 135, 53-65, 1989
A; Title: Structure and tissue-specific developmental expression of a sea urchin aryls A; Reference number: A37362; MuID:89357267
A; Reference number: A37362; MuID:89357267
A; Residues: 1-567 < YAN>
A; Molecule type: mRNA
A; Residues: 1-567 < YAN>
A; C; Superfamily: animal sulfatase
C; Reywords: sulfuric ester hydrolase
C; Reywords: sulfuric ester hydrolase
F; 115/Modified site: 3-oxoalanine (Cys) #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               steryl-sulfatase (EC 3.1.6.2) precursor - human NyAlternate names: arylsulfatase C; steroid sulfatase (STS); steryl-sulfate sulfohydr C;Species: Homo sapiens (man) C;Species: Homo sapiens (man) C;Date: 21-May-1990 #sequence_revision 27-Oct-1995 #text_change 11-May-2000
                                                                                                                                                                                 423 NVSGV---TTHN----LEDHTKLPLIFHLGRDPGERFPLSFASAEYQEALSRITSVVQQHQ 476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               229 AWDYDNGIYSTQMYTQR-VQQIL-----ASHNPTKPIFLYIAYQAVHSPLQAPGRYFE 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LLAQAGEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKTP-TLDKLAAEGVKL-EN 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    116 YYVQPICTPSRSQFITGKYQIHTGL--QHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHM 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         YSGDSVCTPSRSAIVTGRQPIRTGVYGEERIFLPWTTTGLPLYEVTIAEAMKGAGYTTGM 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             281 HYRSIININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNG-----GQPTA 334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SDDFSCTSRRGRYGDNLREMDQAIEQIVTTLVDNDIDDNTVIFFTSDHGPHREYCGE--- 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDE 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      334 GGDANVFRGGKGQSWEGGHRIPYIVYWPGTISPG-VSHEIVTSMDIIATAVNLGGSQLPT 392
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LRSPRVDILHNIDPIYTKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGNPGYSDWVPPQ
                                                                                     376 RLMDRPIFYYRGDTLMA-ATLGQHKAHFWTWTNS------WENFRQG---IDFCPGQ
                                                                                                                                       SFSNLGPNRWHNERITSSTGKSVWLFNITADPYERVDL---SNRYPGIVKKLLRRLSQFN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             174 VGKWHLGFYRKEC-----MPTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          567;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DIQLDGYDIWETISEGLRSPRVDILHNIDPIYTKAKNGSWAAGY 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 13.8%; Score 416.5; DB 2; Best Local Similarity 29.7%; Pred. No. 6.4e-24; Matches 120; Conservative 66; Mismatches 175;
                                                                                                                                                                                                                                                 KTAVPVRYPPKDPRSNPRLN 547
                                                                                                                                                                                                                                                                                                   ----- AQPOLN 487
                                                                                                                                                                                                                                                                                                     EALVP-
                                                                                                                                                                                                                                                                                                     477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28
                                    411
                                                                                                                                                                                                                                                    528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                   a
                                                                                                                                                                                       g
                                                                                                                                                                                                                                                    οy
                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                    ζ
                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Map position: 16q24-16q24
A;Introns: 40/3; 82/1; 107/1; 141/2; 189/2; 211/3; 253/2; 300/1; 334/3; 380/2; 414/3; 45
A;Note: defects in this gene can cause mucopolysaccharidosis type IV A, Morquio disease
C;Function:
                                                                                                                                  ashi, Y.; Orii, T.

Blochem. Blochys. Res. Commun. 181, 677-683, 1991

A;Title: Morquio disease: isolation, characterization and expression of full-length cDNP

A;Reference number: JQ1299; MUID:92095973

A;Recession: JQ1299
                                                                                                                                                                                                                                                                                                                                                                              N.Alternate names: chondroitinase; chondroitinsulfatase; galactose-6-sulfate sulfatase; C;Species: Homo sapiens (man)
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1992 * Resquence_revision 27-Oct-1995 *text_change 08-Dec-2000
C;Accession: JQ1299; PQ0242; I37406
R;Tomatsu, S.; Fukuda, S.; Masue, M.; Sukegawa, K.; Fukao, T.; Yamagishi, A.; Hori, T.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Description: hydrolyzes N-acetylgalactosamine-6-sulfate units in chondroitin sulfate C; Superfamily: animal sulfatase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Molecule type: DNA
A;Residues: 1-522 <RES>
A;Cross-references: EMBL:U06088; NID:g507365; PIDN:AAC51350.1; PID:g618426
A;Note: nucleotide sequence not complete
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TINAHARNAYTPQEIVGGIPDSEQLLPELLKKAGYVSKIVGKWHLG-HRPQFHPLKHGFD 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         202 EANLTQIYLQEALDFIKRQARHHPFFLYWAVDATHAPVYASKPF-----LGTSQRGRYG 255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEG 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DDQGFRDVGYHGSEIK-TPTLDKLAAEGVKLENYY-VQPICTPSRSQFITGKYQIHTGL- 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DAVREIDDSIGKILELLQDLHVADNTFVFFTSDNGAALISAPEQGGSNGPFLCGKQTTFE 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ----QHSIIRPTQ-PNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFD 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EWFG------SPN-CHFGPYDNKARPNIPVYRDWEMVGRYYEEFPINLKTG 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 MAAVVAATRWWQLLL-------VLSAAG---MGASGAPQPPNILLLIM 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24 MLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLAQAGEKLEPSTTSTSQPHLIFILA 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C; Keywords: glycoprotein; sulfuric ester hydrolase
C; Keywords: glycoprotein; sulfuric ester hydrolase
F; 1-26/Domain: signal sequence #status predicted <SIG>
F; 27-522/Product: N-acetylgalactosamine <C-sulfatase #status experimental
F; 79/Modified site: 3-oxoalanine (Cys) #status predicted
F; 204, 423/Binding site: carbohydrate (Asn) (covalent) #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----TQMYTQRVQOILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRSIININRRRYA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 295 AMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNG----GQPTAGGSNWPLRGSKGTYWE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNAA------WDYDNGIYS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 14.0%; Score 422.5; DB 1; Length 522; 25.4%; Pred. No. 2e-24; ative 84; Mismatches 225; Indels 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Status: preliminary; translation not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Gene: GDB:GALNS
A;Cross-references: GDB:129085; OMIM:253000
                                                                                                                                                                                                                                                                    A; Molecule type: mRNA
A; Residues: 1-522 <TOM>
A; Experimental source: placenta
A; Accession: PQ0242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Simi
Matches 142;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

14;

```
S----NRYPGIVKKLLRRLSQFNKT 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         516 TPASEPRFYEILKVMQEAADRHTQT 540
                                                                                                                              Conservative
                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Accession: S69359
                                                                                Best Local Sim
Matches 142;
                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          509
                                                                                                                                                                                                                                                                                      Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A Molecule type: mRNA

A Residues: 1-22, E', 24-456, LRTHPSGRPFSSPPTSTPWFQRIALPHTCASVSGVMSPITTHLYSLIFP', KIPERE

A) Residues: 1-22, E', 24-456, LRTHPSGRPFSSPPTSTPWFQRIALPHTCASVSGVMSPITTHLYSLIFP', KIPERE

A) Note: part of this sequence revised in A57116

A) Note: part of this sequence, including the amino end of the mature protein, was determ

B) Note: part of this sequence, including the amino of the mature protein, was determ

R) Rawano, J.: Kotani, T.; Ohtaki, S.; Minamino, N.; Matsuo, H.; Olnuma, T.; Aikawa, E.

Biochim. Blophys. Acta 997, 199-205, 1989

A) Reference number: 805414; MUID: 89352671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Molecule type: protein
A; Mestdues: 22-43, X', 45-46 <KAM>
A; Residues: 22-43, X', 45-46 <KAM>
B; Dibbelt, L.; Otto, J; Kuss, B.
Biol. Chem. Hoppe-Seyler 370, 847-848, 1989
A; Title: The N-terminal amino-acid sequence of human placental sterylsulfatase.
A; Reference number: $05423; MuID:90074181
A; Reference number: $05423; MuID:90074181
A; Residues: 22-35 <DIB>
A; Residues: 22-35 <DIB
A; Residues: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hyd
C;Accession: A32641; A57116; A25961; S05415; S05423; I52800; I65619
R;Stein, C.; Hille, A.; Seidel, J.; Rijnbout, S.; Waheed, A.; Schmidt, B.; Geuze, H.; vc
J. Biol. Chem. 264, 13865-13872, 1989
A;Title: Cloning and expression of human steroid-sulfatase. Membrane topology, glycosyla
A;Reference number: A32641; MUID:89340479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type: mRNA
A; Residues: 1-22, E', 24-583 < YEN1>
A; Cros. --references: GB:M16505; NID:9338513; PIDN:AAA60596.1; PID:9338514
A; Cros. --references: GB:M16505; NID:9338513; PIDN:AAA60596.1; PID:9338514
R; Yen, P.H.; Allen, E.; Marsh, B.; Mohandas, T.; Wang, N.; Taggart, R.T.; Shapiro, L.J.
Cell 49, 443-454, 1987
A; Title: Cloning and expression of steroid sulfatase cDNA and the frequent occurrence of
A; Reference number: A25961; MUID:87187642
A; Accession: A25961
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Cross-references: GB:J04964; NID:g338564; PIDN:AAA60597.1; PID:g338565
A;Experimental source: BHK-21 cells
A;Note: parts of this sequence were determined by protein sequencing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;MOJECULE type: DNA
A;Residues: 461-583 <YEN4>
A;Cross-references: GB:M23556; NID:g338605; PIDN:AAA60599.1; PID:g338608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Cross-references: GB:M23945; NID:g338604; PIDN:AAA60598.1; PID:g338607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Cross-references: GDB:120393; OMIM:308100
A;Map position: Xp22.32-Xp22.32
A;Note: defects in this gene can cause X-linked ichthyosis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             cited by GenBank
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 134-274 <YEN3>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Accession: 165619
A;Status: translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         unpublished results 1988, cA. Reference number: A57116
A. Accession: A57116
                                                                                                                                                                                                                                                A; Molecule type: mRNA
A; Residues: 1-583 <STE>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Accession: I52800
                                                                                                                                                                                                              Accession: A32641
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Gene: GDB; STS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Genetics
```

```
A,Molecule type: protein
A,Residues: 2-27:95-101;131-155;159-169;188-195;245-263;297-302;340-350;383-388;407-4
A;Residues: 2-27:95-101;131-155;159-169;188-195;245-263;297-302;340-350;383-388;407-4
B;Dierks, T.; Miech, C.; Hummerjohann, J.; Schmidt, B.; Kertesz, M.A.; von Figura, K.
J. Biol. Chem. 273, 25560-25564, 1998
A;Title: Posttranslational formation of formylglycine in prokaryotic sulfatases by mo A;Reference number: A59074; MUID:98421466
A;Contents: annotation; post-translational modification
A;Note: confirmation of 3-oxoalanine active site, referred to as formylglycine
C;Genetics:
                                                                           22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             þ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C; Species: Pseudomonas aeruginosa
C; Species: Pseudomonas aeruginosa
C; Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 17-Mar-2000
C; Accession: S69336; S69359
R; Beil, S.; Kehrli, H.; James, P.; Staudenmann, W.; Cook, A.M.; Leisinger, T
Eur. J. Biochem. 229, 385-394, 1995
A; Title: Purification and characterization of the arylsulfatase synthesized 1
A; Reference number: S69336; MUID:95262702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A,Molecule type: DNA
A,Residues: 1-533 <BEI>
A;Cross-references: EMBL:248540; NID:92440146; PIDN:CAA88421.1; PID:9695684
A;Experimental source: strain PA01
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            306 GV-----YGDAVEEMDWSVGQILNLLDELRLANDTLIYFTSDQGAHVEEVSSKGEIHGG 359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     454
                                                                                                                                                                                                                                                                                                                           125 SRSQFITGKYQIHTGLQH-----SIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHL 179
                                                                                                                                                                                                                                                                                                                                                                 138 GM---SCHSKTDFCHHPLHHGFNYFYGISLTNLRDCKPGEGSVFTTGFKRLVFLPLQIVG 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              195 VILLTLAALNCEGLEHVPEGVFFSLLFLAALILTEFLGFLHYFRPLNCFM--MRNYEIIQ 252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  225 NDNAAWDYDNGIYSTQMYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRS 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         337 SNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDI 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    464
                                                                                                                                                               67 EPSTTSTSQPHLIFILADDQGFRDVGYHGSE-IKTPTLDKLAAEGVKL-ENYYVQPICTP 124
                                                                                                                                                                                                     465 KAFFFTP-----NFNPVGSNGCFATHVCFCFGSYVTHHDPPLLFDISKDPRERNPL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q---PMSYDN----LTQRLTVEAAQFI-QRNTETPFLLVLSYLHVHTALFSSKDFAGKSQH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           || :|| || || || || | : | : || || || 360 SNGIYKGGKANNWEGGIRVPGILRWPRVIQAGQKIDEPTSNMDIFPTVAKLAGAPLPEDR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KLLTGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGKSV-----WLFNITADPYERVDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         397 QLDGYDIWETISEGLRSPRVD--ILHNIDPIYTKAKNGSWAAGYGIWNTAIQSAIRVQHW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    W----WHPONSTSI---W
    Length 583;
13.7%; Score 413; DB 1; Length 583
25.1%; Pred. No. 1.2e-23;
Live 78; Mismatches 201; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             285 IININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNGGQ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      180 GFYRKEC------MPTRRGFDTFFGSLL------GSGDYYT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       arylsulfatase (EC 3.1.6.1) - Pseudomonas aeruginosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C; Superfamily: animal sulfatase
```

ð g ö g ð g ò a ò 셤 ò 셤 ò 셤 ô g ç g ò g

```
arylsulfatase D (EC 3.1.6.-) - human C;Species: Homo sapiens (man)
C;Species: Homo sapiens (man)
C;Species: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 22-Jun-1999
C;Accession: 137186
R;Franco, B.; Meroni, G.; Parenti, G.; Levilliers, J.; Bernard, L.; Gebbia, M.; Cox, Cell 81, 15-25, 1995
A;Title: A cluster of sulfatase genes on Xp22.3: mutations in chondrodysplasia puncta A;Reference number: A56217; MUID:95236447
A;Steference number: A56217; MUID:95236447
A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    27;
                                                                     25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 265 P-MC----FQR------TTPLILQEVASFL-KRNKHGPFLLFVSFLHVHIPLIT 306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A;Molecule type: mRNA
A;Residues: 1-593 <RES>
A;Cross-references: EMBL:X83572; NID:9791001; PIDN:CAA58555.1; PID:9791002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      384
                                                                                                                                  66 LEPSTT---STSQPHLIFILADDQGFRDVGYHGSE-IKTPTLDKLAAEGVKL-ENYYVQP 120
                                                                                                                                                                                                                                                                                                                                                                                                                444 AIQSAIRVQHWKLLTGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGKSV-----WLFN 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                          359 ENQLGNTQYGGWNGIYKGGKGMGGWEGGIRVPGIFRWPGVLPAGRVIGEPTSLMDVFPTV
                                                                                                                                                                  25 LAPSASSDISASRPNILLLMADDLGIGDIGCYGNNTMRTPNIDRLAEDGVKLTQHISAAS
                                                                                                                                                                                                                                                    121 ICTPSRSQFITGKYQIHTGLQHSI-IRPTQ----PNCLPLDNATLPQKLKEVGYSTHMVG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --GDYYTHYKC-----DS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               205 ALVALTLVAGKLTHLIPVSWMPVIWSALSAVLLLASSYFVGALIVHADCFLMRNHTITEQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ----GQPTAGGSNWPLRGSKGT-YWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            385 ISLAEGQIDEDIQLDGYDIWE-TISEGLRSPRVDILHNIDPIYTKAKNGSWAAGYGIWNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        469 RDRGTM----WKV------HFVTP-VFQPEGAGACYGRKVCPCFGEKVVHHDPPLLFD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PGRYFEHY --- RSIININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 13.2%; Score 399; DB 2; Length 593; Best Local Similarity 24.4%; Pred. No. 1.5e-22; Matches 144; Conservative 87; Mismatches 191; Indels 168;
                                                                     Indels 161;
         Length 589;
                                                                                                                                                                                                                                                                                                                                                                             KWHLGFYRKECM-----PTRRGFDTFFG---SLLGS-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; Pred. No. 3.3e-23; 92; Mismatches 188;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       498 ITADPYERVDLSNRYPGIVKKLLRRLSQ----FNKTAVPV 533
         DB 2;
         Score 407.5;
Pred. No. 3.36
      13.5%; 24.0%;
Query Match 13.59
Best Local Similarity 24.09
Matches 139; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Gene: GDB: ARSD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                         145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 275
                                                                                                                                                                                                                                                                                                                                                                             176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      330
                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ρp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                  ð
                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              arylaufatase E (EC 3.1.6.-) - human (Species: Homo sapiens (man) (Species: Homo sapiens (man) (Species: Homo sapiens (man) (Species: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 22-Jun-1999 (Stacession: 137187 (Stanco, B.; Meroni, G.; Parenti, G.; Levilliers, J.; Bernard, L.; Gebbia, M.; Cox, L.; Riffanco, B.; Meroni, G.; Parenti, G.; Levilliers, J.; Bernard, L.; Gebbia, M.; Cox, L.; A.Title: A cluster of sulfatase genes on Xp22.3: mutations in chondrodysplasia punctata A. Reference number: A56217; MUD:95236447 (A. Reference number: B56217; MUD:95236447 (A. Reference number: B782187 (A. Reference number: B782187 (A. Reference number: B788187) (A. Reference number
                                                                                                                                                                                           20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---LYEN 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    227 RYDAGPEALRQERLARLKELGLVEADVEAHPVLALTREWEALEDEERAKSARAMEVYAAM 286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VERMDWNIGRVVDYLRROGELDNTFVLFMSDNA - - PKAPCWRRSRIGRTCWLLDRHYDNS 344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         73 TSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITG 132
                                                                                                                                                                                                                                                                                                                                                                                                                377 ITDWYPTLISLAEGQIDEDIQLDGYDIWETISE---GLRSPRVDILHNIDPIYTKAKNGS 433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WAAGYGIWNTAIQSAIRVQHWKLLTGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGKSV 493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ----PAT 483
                                                                                                                                                                                                                                                                                                                                                                       133 KYQIHTGLQHSII--------RPTQPNCLPLDNATLPQKLKEVGYSTHMVGKW 177
                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                    2 SKRPNFLVIVADDLGFSDIGAFGGEIATPNLDALAIAGLRLTDFHTASTCSPTRSMLL-- 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DNAAWD-YDNGIYSTQMYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LSCLDEAINNVTLALKTYGFYNNSIIIYSSDNGGQPTAG-----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------GSNW-----PLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VMDVTPTLLDLA-GVRHPGKRWRGREIAEPRGRSWLGWLSGETEAAHDENTV-----
                                                                                                                                                                                        Indels 179;
                                                                                                                           Length 533;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HLGFYRKECMPTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCGYD-----
C;Keywords: sulfuric ester hydrolase
F;51/Modified site: 3-oxoalanine (Cys) #status experimental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             --TG---WELFGMRAIRQGDWKAV-------YLPARW-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A.Gene: GDB:ARSE; CDPX; CDPX1
A.Cross-references: GDB:555743; OMIM:302950
A.Map position: Xp22.3-Xp22.3
C.Superfamily: animal sulfatase
C.Keywords: sulfuric ester hydrolase
F:86/Modified site: 3-oxoalanine (Cys) #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         494 W-LFNITADPYERVDLSNRYPGIVKKLLRRLSQF-NKTAV 531
                                                                                                                                                                                        67; Mismatches 191;
                                                                                                                               DB 2;
                                                                                                                                                           Pred. No. 1.2e-23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                484 WQLYDLARDPGEIHDLADSQPGKLAELIEHWKRYVSETGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ---IININE---
                                                                                                                           13.7%; Score 412.5; 24.7%; Pred. No. 1.2
                                                                                                                                                        Best Local Similarity 24.7%
Matches 143; Conservative
                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        405
                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          285
```

OY 114 ENYYOQPICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYS 170	Db 341 ENGQLDHTLICTSDNGPEAEVPHGRTPERGAGSTWEGGVRVFFVWKG 393  Qy 369 TVCKEPVHITDWYPTLISLAEQDIDEDIQLDCYDIWETISEGLRSPRVDILHNI 422  E :	RESULT 14 E86066 arylsulfatase [imported] - Escherichia col1 (strain O157:H7, substrain EDL933) C; Species: Escherichia col1 C; Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001 C; Accession: E86066 R; Perna, N.T.; Plunkett III, G; Burland, V; Mau, B; Glasner, J.D.; Rose, D.J.; May iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousls, K.; Apoda Nature 409, 529-533, 2001 A; Title: Genome sequence of enterchemorrhagic Escherichia col1 0157:H7. A; Reference number: A65480; MUD:21074935; PMID:11206551 A; Status: preliminary A; Status: preliminary A; Status: GB:AE005174; NID:g12518665; PIDN:AAG58993.1; GSPDB:GN00145; UWGP: A; Experimental source: strain 0157:H7, substrain EDL933	A; Gene: aslA         Query Match       13.2%; Score 39%; DB 2; Length 551;         Best Local Similarity 24.6%; Pred. No. 1.68-22;         Matches 148; Conservative 97; Mismatches 204; Indels 152; Gaps 29;         Qy 2 ARGAGHPPPPSPOACVCPGKMLAMGLAGFWILCLITYGYLSWGALEEEEEGA 57
67 EPSTTSTSQPHLIFILADDOGFRDVGYHGSE-IKTPTLDKLAABGUKL-ENYYQQPICTP [	281 HYRSIININRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNGGQPTA	0509952) K.; Han, C.G. H.	Cross-references: GB:BA000007; PIDN:BAB38154.1; PID:g13364207; GSPDB:GN00154 Experimental source: strain 0157:H7, substrain RIMD 0509952 Genetics: Gene: ECs4731  Duery Match  13.2%; Score 39%; DB 2; Length 551; Best Local Similarity 24.6%; Pred. No. 1.6e-22; Matches 148; Conservative 97; Mismatches 204; Indels 152; Gaps 29;  2 APRGCAGHPPPPSPQACVCRKMLAMARCHMICLLITYGYLSWGOALEEEEEGA 57 2 APRGCAGHPPPPSPQACVCRKMLAMAPUWQHPA

. . .

```
:| : : | | | | : : | | | 323
                                                                                                                                                                                                                                                                                                                                                                                                      120 PICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHL 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----LISLAEGQIDEDIQLDGYDIWETISEGLR-SPRVDILHNIDPIYTKAKNGSWAA 436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   437 GYGIWNTAIQ-SAIRVQHWKLL---TGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGKS 492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       493 VWLFNITADPYERVDLSNRY-------PGIVKKLLRRLSQFNKTAVPVRYPP 537
                                                                                                                                     GFYRKECMPTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYST
                                                                                                                                                                                                QMYTQRVQQILASH--NPTKPI-------FLYIAYQAVHSPLQAPGRYFEHYRSI-
                                                                                                                                                                                                                                            204 RFGPRGVIKATADGKIEDTGPLSVKRMETVDQEFLEAALDFIDRKAKGDAPWFCYFNSTR
                                                                                                                                                                                                                                                                                       --ININRRRYAAMLSCL------DEAINNVTLALKTYGFYNNSIIIYSSDNGGQPT
                                                                                                                                                                                                                                                                                                                                                                        334 A--GGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DIVAKCLIGYAAGAKTFKVHLDGYNLMPFLSGSSNDAPRRDFLY ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Search completed: July 23, 2002, 19:04:13
Job time: 6549 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               467
                                                                                                          180
                                                                                                                                                                                                240
                                                                                                                                                                                                                                                                                                                                                                                                                                                               384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       384
                                                                                                                                                                                                                                                                                       286
                                                        Op
                                                                                                                                          g
                                                                                                                                                                                              ò
                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                           ò
                                                                                                          ò
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: DNA
A; Residues: 1-537 < KUR>
A; Experimental source: strain 1021, megaplasmid psymb
B; Galibert, F.; Finan, T.W.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Cabin, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001
A; Aluthors: Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Weils, D.H.; Wong, K.; Yeh, K.A; Title: The composite genome of the legume symbiont Sinorhizoblum meliloti.
A; Reference number: A96039; MUID: 21368234; PMID: 11474104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        probable arylsulfatase (EC 3.1.6.1) [imported] - Sinorhizoblum meliloti (strain 1021) me C; Species: Sinorhizoblum meliloti
C; Species: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
C; Accession: F95993
R; Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A; Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endchance number: A95842; MUID:21396508; PMID:11481431
A; Accession: F95993
A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     22;
                                                                                                                                                                                                                                                368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                128 TSAYSQPSSSPTRATILIGQYSIH----HGILMPPMYGQPGGLQ-GLTTLPQLLHDQGYV 182
                                                                                                                                                                       230 ALSPDRSEYIKOLPFSKDDVHAVRGGEQOAIADITPKYMEDLDQRWMEYGVKFLDKMAKS 289
                                                                                                                                                                                                                                                                                                                                                                                           369 TVCKEP-----VHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNI 422
                                                                                                                                                                                                                                                                                                                                                                                                                    NRWHNERITSSTGKSVWLFNITADPYERVDLSNRYPGIVKKLLRRLSQFNKTAVPVRYPP 537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       489 QGGFTGTVMQTAGSSV--FNLYTDPQESDSIGVRHIPMGVPLQTEMHAYME--ILKKYPP 544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KLEPSTTSTS--QPHLIFILADDQG-FRDVGYHGSEI--KTPTLDKLAAEGVKLENYYVQ 119
                                                                 183 TQAIGKWHWG-ENKESOPONVGFDDFRGFNSV---SDMYTEWR----DVHVNPFV
                                                                                                                                                                                                                                                                                                                                  THMVGKWHLGFYRKECMPTRRGFDTF - - FGSLLGSGDYYTHYKCDSPGMCGYDLYENDNA
                                                                                                                                                                                                                                                                                                         312 KTYGFYNNSIIIYSSDNGGQ---PTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | :: | : | : | : | | : | 438 DQTSFFLGTNGQSNRKAEHYFL--NGKLAAVRMDEFK------YHVLIQQPYAYTQSGY
                                                                                                                                                                                                                   TKPIFLY1AYQAVHSPLQAPGRYFEHYRSI----ININRRRYAAMLSCLDEAINNVTLAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DP----IYTKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGNPGYSDWV-PPQSFSNLGP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels 127;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13.1%; Score 395.5; DB 2; 25.8%; Pred. No. 2.4e-22; attive 78; Mismatches 195;
                                                                                                                               AWDYDNGIYSTQMYTQR-----VQQILASHNP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Genome: plasmid
C;Keywords: sulfuric ester hydrolase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Gene: aslAl; SMb20915
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Contents: annotation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 139;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         538 K 538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            65
                                           171
                                                                                                                            229
                                                                                                                                                                                                                   256
                                                                                                                                                                                                                                                              290
                                                                                                                                                                                                                                                                                                                                                  341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  545
  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                           ò
                                                                                 셤
                                                                                                                            à
                                                                                                                                                                 g
                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
```

```
Humon sapiens

Bukaryota: Metazoa: Chordata; Craniata: Vertebrata; Euteleostomi; Mamalia: Eutheria: Primates; Catarrhini; Hominidae; Homo.

E 1 (bases 1 to 2253)
S Glucksmann.M.A., Williamson,M., Rudolph-Owen,L.A. and Tsai,F.Y.
Human sulfatases
L patent: WO 0155411-A 8 02-AUG-2001;
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
irce //db_xref="Homo sapiens"
//db_xref="Homo sapiens"
//db_xref="Homo sapiens"
//db_xref="Laxon:9606"
//db_xref="Laxon:9606"
//codon_start=1
                                                                                                                                                                                M32373 Human aryls
J05225 Human aryls
D49434 Ratt ARSB mR
AC098604 Rattus no
AC105383 Homo sapi
X22096 M.musculus
AF109925 Helix pom
BC020108 Mus muscu
                                                                                                                                                                                                                                                                                                  ACO 100 S Rattus no X72739 H.sapiens 9 ACO 99468 Homo sapi ACO 91369 Rattus no ACO 13689 Homo sapi ACO 87276 H.sapiens 9 ACO 2037 Homo sapi ACO 2037 Homo sapi ACO 2037 Homo sapi ACO 20310 Homo sapi ACO 20310 Homo sapi ACO 20310 Homo sapi ACO 20310 Homo sapi ACO 69033 Homo sapi ACO 69033 Homo sapi ACO 20310 Homo sapi ACO 203110 Homo sapi ACO 20312 Drosophil ACO 10352 Drosophil ACO 10352 Drosophil ACO 10352 Drosophil ACO 10502 Drosophil ACO 10502 Drosophil ACO 10502 Drosophil
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PAT 30-AUG-2001
                                                                                                    AC105414 Homo sapi
AX206965 Sequence
AX206972 Sequence
AC011372 Homo sapi
AF109924 Helix pom
S48472 arylsulfata
                                                                                          Homo sapi
Homo sapi
                           Homo sapi
Homo sapi
                                                                 Homo sapi
Homo sapi
                                                     Mus muscu
  Sequence
            AX206973 3
AC104779 1
AC021342 1
AC091322 1
AK027201 1
AC013692 1
AC021342 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 8 from Patent WO0155411.
AX206967.1 GI:15394724
                                                                                                                                                        AF109924
S48472
HUMARSBX
HUMASFB
RATARSB
AC098604
AC105383
                                                                                                                                                                                                                                                                              BC020108
                                                                                                                                                                                                                                                                                                                                              AC091369
AC013689
AC008787
                                                                                                                                                                                                                                                                                                                                                                                               AC020937
AC022110
AC008787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AC022110
166494
                                                                                                                                                                                                                                                                                                      AC109105
HSARYB5
                                                                                                                                                                                                                                                                                                                                                                                                                                          HSARYB1
AC069033
               AX206973
AC104779
                                                                                                      AC105414
AX206965
                                                                                                                                                                                                                                                    MMASLS
AF109925
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AE003522
AC014974
                                                                                                                                                                                                                                                                                                                                  AC099485
                                                                                                                                                                                                                                                                                           AY071072
                                                    AC091322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 S57777
HSARYB4
                                                                                                                                                                                                                                                                                                                                                                                      HSARYB2
                                                                                                                                                                        400122
                                                                                                                                                                                                                                                  33
                                                                                                                                                                                                                                                                              10
                                                                                                                                                                                                                                                                                                                                                                                                                                                     3.4 187592
3.3 766
3.7 119582
2.6 7218
2.5 32231
2.5 179649
2.5 179649
2.4 158520
 114624
145718
176356
164478
271
                                                                                                                                                                                                                                                                                                                                                                                                                         164478
1926
187592
766
                                                                                                                                                                                                                                                                            1428
1777
16052
329
                                                                                                                                                                                                                                                                                                                                                                                                 94161
119582
 18777533333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            human.
17234.2
1781.1
11111.1
11111.1
1081.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
1090.6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                55.6
55.6
55.4
55.4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
AX206967
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
  LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             O O
                            000
                                                                            υ
                                                                                                      O
                                                                                                                                                                                                                                                                                                                                                             υυ
                                                                                     July 23, 2002, 15:04:54; Search time 4089.32 Seconds (without alignments) 11595.946 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                     1 cacgogtecgeceaegegte.....tgecaectggtgeegaatte 2266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Description
                                                                                                                                                                                                                                                                3595312
            version 4.5
- 2000 Compugen Ltd.
                                                                                                                                                                                                                                    1797656 seqs, 10463268293 residues
                                                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUMMARIES
                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                               nucleic search, using sw model
                                                                                                                                                                                              IDENTITY_NUC Gapop 10.0 , Gapext 1.0
            GenCore
Copyright (c) 1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              em_htg_inv:*
em_htg_other:*
em_htgo_inv:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                       Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       em_vi:*
em_htg_hum:*
                                                                                                                                          US-09-495-823-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gb_ro:*
gb_sts:*
gb_sy:*
gb_un:*
em_bu:*
em_fun:*
em_fun:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              em_ov:*
em_pat:*
                                                                                                                                                                                                                                                                                                                                                                                               9b_ba:*
9b_in:*
9b_om:*
9b_om:*
9b_ow:*
9b_pat:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     em_ph:*
em_pl:*
em_ro:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       em_mu:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             em_sts:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 em_or:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            em_un:*
                                                                                                                                                                                                                                                                                                                                                                                    GenEmbl:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score
                                                                                                                                                         score:
                                                                                                                                                                                              Scoring table:
                                                               nucleic
                                                                                                                                                                                                                                      Searched:
                                                                                                                                                                     Sequence:
                                                                                                                                                                                                                                                                                                                                                                                    Database
                                                                                                                                           Title:
Perfect :
                                                                                       Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š.
```

```
/protein_id="Cac60191.1"
/db_xref="G1.15394725"
/translation="MaprgcAghPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYL
xref="G1.15394725"
/translation="MaprgcAghPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYL
SWGQALEBEEGALLAQAGEKLEPSTTSTSQPHLIFILADDGGFRDVGYHGSETKTPT
LDKLAAEGVKLENYYVQPTCTPSRSQFITGKYQIHGCLQHSITRPTQPNCLPLDNATL
PQKLKEVGYSTHWVGKWHLGFYRECMPTRRCEDTFPGSLLGSGDYTHYKTOSPEGMC
GYDLYENDNAAWDYDNGIYSTYMYTQRYQOILASHNPTKPIFEYIAYQAVHSPLQAPG
RYFEHYRSIININRRYAAMLSCLDEAINNYTLALKTYGFYNNSIIIYSSDNGQQPTA
GGSNWPPROSKGTYMSGGIRAYGFYBPLLAKNGTYCKELVHTTNWYPILSLAEGQI
DEDIQLOGYDIWETISEGLASPRVDILHNIDPIYTRAKNGSVAAAGYGTWNTAAQSAIR
VQHWKLLTGNPGSDWYPPOSFSNLGPNRWHNERITLSTGKSVWLFNITADPYERVDL
SNRYYGIGYKKLLRAGSNYTAPVDRYPPRDPRSNPRLNGGVWGPWYREETKKKKPSK
NQABKKOKKSKKKKKKQOKAVSGSTCHSGYTCG"

NGABKQKKSKKKKKKQOKAVSGSTCHSGYTCG"
                                                                                                                                                  ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                           900
                                                                                                                                                                                                                                                         240
                                                                                                                                                                                                                                                                      240
                                                                                                                                                                                                                                                                                                                                                                                                                                            009
                                                                                                                                                                                             120
                                                                                                                                                                                                            120
                                                                                                                                                                                                                            180
                                                                                                                                                                                                                                         180
                                                                                                                                                                                                                                                                                       300
                                                                                                                                                                                                                                                                                                     300
                                                                                                                                                                                                                                                                                                                     360
                                                                                                                                                                                                                                                                                                                                   360
                                                                                                                                                                                                                                                                                                                                                   420
                                                                                                                                                                                                                                                                                                                                                                 420
                                                                                                                                                                                                                                                                                                                                                                                480
                                                                                                                                                                                                                                                                                                                                                                                                480
                                                                                                                                                                                                                                                                                                                                                                                                               540
                                                                                                                                                                                                                                                                                                                                                                                                                             540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         720
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        780
                                                                                                                                                 Gaps
                                                                                                                                                               9
                                                                                                                                                                          9
                                                                                                                                                              agccctccttgttcttccggagtcccatccattaagccatcacttctggaagattaaagt
                                                                                                                                                                                                                                                                                                                                                                                                                                            tctggatcctctgcctcacttatggttacctgtcctgggggccaggccttagaagagg
                                                                                                                                                1;
                                                                                                                                  Length
                                                                                                                                                Indels
                                                                                                                                  9
                                                                                                                                                4;
                                                                                                                                   DB
                                                                                                                                 ore 2234.2;
ed. No. 0;
Mismatches
                                                                                                                                  Score Srore Pred. 1
                                                                                                                                 Query Match
Best Local Similarity 99.7%;
Matches 2247; Conservative
                                                                                                    624
                                                                                                   BASE COUNT
ORIGIN
                                                                                                                                                                                             61
                                                                                                                                                                                                            61
                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                       241
                                                                                                                                                                                                                                                                                                     241
                                                                                                                                                                                                                                                                                                                    301
                                                                                                                                                                                                                                                                                                                                   301
                                                                                                                                                                                                                                                                                                                                                  361
                                                                                                                                                                                                                                                                                                                                                                 361
                                                                                                                                                                                                                                                                                                                                                                               421
                                                                                                                                                                                                                                                                                                                                                                                              421
                                                                                                                                                                                                                                                                                                                                                                                                             481
                                                                                                                                                                                                                                                                                                                                                                                                                             481
                                                                                                                                                                                                                                                                                                                                                                                                                                            541
                                                                                                                                                                                                                                                                                                                                                                                                                                                           541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    721
                                                                                                                                                                            QQ
                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                                  oy
Op
                                                                                                                                                               ò
                                                                                                                                                                                            ŏ
                                                                                                                                                                                                          ద
                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                     oy
Og
                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                             ద
                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                           요
                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                         οp
                                                                                                                                                                                                                                                                                                                                                                                                                                         ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
```

1440 1020 1020 1140 1140 1260 1320 1380 1440 1560 1560 1620 1620 1680 1680 1740 1740 840 840 900 960 960 tatggcttttcaacatcacagccgacccatatgagaggtggacctatctaacaggtatc acagtcctgggatgtgtgtgtgtatgacttgtatgaaaacgacaatgctgcctgggactatg cggcaggaggagtaactggcctctcagaggtagcaaaggaacatattgggaaggaggga gcaacctgggaccgaaccggtggcacaatgaacggatcacctcgtcaactggcaaagtg caggaatcgtgaagaagctcctacggaggctctcacagttcaacaaaaactgcagtgccgg 781 841 1021 1021 1081 1081 1141 1141 1201 1201 1261 1261 1321 1321 1381 1381 1441 1441 1501 1561 1621 1681 1681 1741 1741 1801 1801 1861 901 961 1501 1561 1621 841 901 961 781 셤 ò a οy g ò D δ g δý Ω δ g δy qq δ g á Op ò Op ò Ω ò g οy g ò qq ð qq δ g ŏ QQ à

	y 624 cctactcttgacaagctcgctgccgaaggagttaaactggagaactactatgtccagcct 683 	y 684 atttgcacaccatccaggagtcagtttattactggaaagtatcagatacacaccggactt 743 	y 744 caacattctatcataagacctacccaacccaactgtttacctctggacaatgccaccta 803	Oy 804 cctcagaaactgaaggagttggatattcaacgcatatggtcggaaaatggcacttgggt 863 	Oy 864 ttttacagaaaagaatgcatgcccaccagaagagatttgatacctttttggttccctt 923 	Oy 924 ttgggaagtgggattactatacacactacaatgtgacagtcctgggatgtgtgggtat 983 	Oy 984 gacttgtatgaaaacgacaatgctgcctgggactatgacaatggcatatactccacacag 1043	Oy 1044 atgtacactcagagagtacagcaaatcttagcttcccataaccccacaaagcctatatt 1103 	Oy 1104 ttatattycctatcaagctyttcattcaccactycaagctcctygcaggtatttcgaa 1163 	Oy 1164 cactaccgatccattatcaacataaacaggaggagatatgctgccatgcttcctgctta 1223 	Oy 1224 gatgaagcaatcaacaaggacattggctctaaagacttatggttctataacaacagc 1283 	Oy 1284 attatcatttactcttcagataatggtggccagcctacggcaggaggagtaactggcct 1343	Oy 1344 ctcagaggtagcaaaggaacatattgggaaggaggatccgggctgtaggctttgtgcat 1403 	Oy 1404 agcccacttctgaaaaacaagggaacagtgtgtaaggaacctgtgcacatcactgactg	Oy 1464 taccccactctcatttcactggctgaaggacagattgatgatgagacattcaactagatggc 1523	Oy 1524 tatgatatctgggagaccataagtgaggtcttcgctcaccccgagtagatattttgcat 1583	Qy     1584 aacattgaccccatatacaccaaggcaaaaatggctcctgggcagcagctatggatc 1643       Db     111111111111111111111111111111111111	Oy 1644 tggaacactgcaatccagtcagccatcagagtgcagcactggaaattgcttacaggaaat 1703
qq (	Oy Op	QQ Op	O O O O O O O O O O O O O O O O O O O			QQ Op		7001 QQ	i,		QQ Op	о a	oy on op		da da	· ·	O 0	
Db 1861 CAGGAATCGTGAAGAGCTCCTACGGAGCTCTCACAGTTCAACAAAACTGCAGTGCCGG 1920	Oy 1921 tcaggtatcccccaaagaccccagaagtaaccctaggctcaatggagggtctagggac 1980	Oy 1981 catggtatagagagaaaccaagaaaagagccaagcaaaaatcaggctgagaaaagc 2040	OY 2041 aaaagaaaagc-aaaaaagaagaagaaacagcagaaagcagtctcaggttcaacttgcc 2099	Oy 2100 attcaggtgttacttgtggataagcacaaatatttcctgtttggttaaactttaatcagt 2159	Oy 2160 tettatettteatetgttteetaggtaaaceageaaatttggetegataatategetgge 2219	<pre>Qy 2220 ctaagcgtcaggcttgtttcatgctgtgccac 2252                                      </pre>		DEFINITION Sequence 14 from Patent W00155411. ACCESSION AX206973 USESSION AX206973.1 G1:15394729	human. SM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eu	. 🚅 🧏	Millenniu ce	### A 434 COUNT 534 A 434 C 423 G 409 t CARGIN	Query Match 78.6%; Score 1781.6; DB 6; Length 1800; Best Local Similarity 99.7%; Pred. No. 0; Matches 1795; Conservative 0. Mismatches 4. Indale 1. Care	atggctcccaggggctgtgcggggcatccgccttctccacaggcctgtgtctgt 3	384 cctggaaagatgctagcaatgggggcgtggcaggattctggatcctctgctcact 4	444 tatggttacctgtcctggggccaggccttagaagaggaggaagaagagggccttactgct 5   11   11   11   11   11   11   11	504 caagctggagagaaactagagcccagcacaacttccacctcccagccccatttc 56	564 atcctagcggatgatcagggatttagagatgtgggttaccacggatctgagattaaaaca 62

1;

```
Query Match
           COUNT
                                                                                                                                                                                                                                                                                                            1079
                                                                                                                                                                                                                                                                                                                                                                 9779
                                                                                                                                                                                                                                                                                                                                                                                    1199
                                                                                                                                                                                                                                                                                                                                                                                                                        1259
                                                                                                                                                                                                                                                                                                                                                                                                                                          9659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9539
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1619
                                                                                                                                                                                                                                   959
                                                                                                                                                                                                                                                     9959
                                                                                                                                                                                                                                                                                                                                                1139
                                                                                                                                                                                                                                                                                                                                                                                                      9719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9299
          BASE CC
ORIGIN
                                                                                    ò
                                                                                                     g
                                                                                                                                        g
                                                                                                                                                                            g
                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ω
                                                                                                                        ò
                                                                                                                                                           ò
                                                                                                                                                                                               à
                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                      Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                           PRI 09-FEB-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Louis,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Louis
                                                                                                                                                                                                                                                                                                                                                                          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 152861)
Waterston, R.H.
                                                        1823
                                                                         1500
                                                                                           1943
                                                                                                                                                 1620
                                                                                                                                                                  2062
                                                                                                                                                                                                                          1740
                                                                                                                                                                                                                                            TGGAACACTGCAATCCAGTCAGCCATCAGAGTGCAGCACTGGAAATTGCTTACAGGAAAT 1380
                                                                                                                                                                                                                                                                                                            DNA linear PRI 09-FEE RP11-26P13, complete sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                Direct Submission
Submitted (21-DEC-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. MO 63108, USA
3 (bases 1 to 152861)
Waterston, R. H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Direct Submission
Submitted (09-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St.
MMO 63108, USA
On Feb 9, 2002 this sequence version replaced gi:18482303.
                  aaaaaggaagccaagcaaaaatcaggctgagaaaagcaaaagaaaagc-aaaaaagaag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Center: Washington University Genome Sequencing Center Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Information
                                                                                                                                                                                                                                                                                                                                                                                                              clone
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1. .152861
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
                                                                                                                                                                                                                                                                                                        152861 bp
Homo sapiens chromosome 4 clone
AC104779
AC104779.4 GI:18642899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Center project name: H_NH0026P13
                                                                                                                                                                                                                                                                                                                                                                                                            The sequence of Homo sapiens
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ----- Project
                                                                                                                                                                                                                                                                                                                                                                                                                               2 (bases 1 to 152861)
Waterston, R.H.
                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens
                                                                                                                                                                                                                                                                                                                                                         human.
                                                                                                                                                                                                                                                                                       RESULT 3
AC104779/C
LOCUS
                                                                                                                                                                                                                                                                                                                   DEFINITION
ACCESSION
                                                                                                                                                                                                                                                                                                                                               KEYWORDS
SOURCE
ORGANISM
                                                                                                            1501
                                                                                                                                                                  1944
                                                                                                                                                                                    1621
                                                                                                                                                                                                                                           2063
                  1704
                                                       1764
                                                                        1441
                                                                                                                                                 1561
                                                                                                                                                                                                       2004
                                                                                                                                                                                                                                                                                                                                                                                                    AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                        AUTHORS
TITLE
JOURNAL
1321
                                    1381
                                                                                           1824
                                                                                                                              1884
                                                                                                                                                                                                                         1681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                      VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMMENT
a
                  ö
                                   a
                                                                     g
                                                                                          ŏ
                                                                                                           Q
                                                                                                                              ò
                                                                                                                                              g
                                                                                                                                                                  ò
                                                                                                                                                                                    g
                                                                                                                                                                                                     ò
                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                          q
                                                       õ
```

```
10140
                                                                                                                                                                                                                                                                                                                                             1138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9240
                                                                                                                                                                                                                                                                                                                                                                                     1198
                                                                                                                                                                                                                    958
                                                                        Gaps
                                                                                                                                 719 aaagtatcagatacacaccggacttcaacattctatcataagacctacccaacccaactg
                                                                                                      gacttatggtttctataacaacagcattatcatttactcttcagataatggtggccagcc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9479 GGAACTTGTGCACATCACTGACTGGTACCCCACTCATTTCACTGGCTGAAGGACAGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gcaagctcctggcaggtatttcgaacactaccgatccattatcaacataaacaggaggag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GACTTATGGTTTCTATAACAACAGCATTATCATTTACTCTTCAGATAATGGTGGCCAGCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ggaacctgtgcacatcactgactggtaccccactctcatttcactggctgaaggacagat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ctcaccccgagtagatattttgcataacattgaccccatatacaccaaggcaaaaatgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CTCACCCCGAGTAGATATTTGCATAACATTGACCCCATATACACCCAAGGGAAAAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       tgatgaggacattcaactagatggctatgatatctgggagaccataagtgagggtcttcg
                                                                       ij
                                                  Length 152861;
                                                                       Indels
                                                                       5,
                                                  6
                                                   DΒ
                                         Score 1515; F
                                                                       0; Mismatches
          47805
/clone="RP11-26P13"
48870 a 27681 c 28505 g
                                                  66.9%;
99.6%;
                                                            Best_Local Similarity 99.6
Matches 1529; Conservative
```

Wed Jul

ò

ò 셤 g

õ g ŏ g à g

ò

ò g g

õ

ð

```
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Center: Whitehead Institute/ MIT Center for Genome Research
                                               Zimmer,A. and Zody,M.

Direct Submission
Submitsolon
Submitsolon
Submitsolon
Submitsolon
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 9, 2000 this sequence version replaced gi:6899755.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           * NOTE: This is a 'working draft' sequence. It currently consists of 14 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Center project name: 14477

Center clone name: 24.1.3

Center clone name: 24.1.3

Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731

Consensus quality: 156719 bases at least Q40

Consensus quality: 160779 bases at least Q30

Consensus quality: 160377 bases at least Q20

Insert size: 170000; agarose-fp

Insert size: 170000; agarose-fp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Quality coverage: 4.4 in Q20 bases; agarose-fp Quality coverage: 4.6 in Q20 bases; sum-of-contigs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Contact: sequence_submissions@genome.wi.mit.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25209 25308: gap of 100 bp 25309 33458: contig of 8150 bp in length 33459 3358: gap of 100 bp 100 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      122717 122816: gap of 100 bp 122817 149492: contig of 26676 bp in length 149493 149592: gap of 100 bp 149593 165147: contig of 15555 bp in length. Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60566 60665: gap of 100 bp 60666 92690: contig of 32025 bp in length 92691 92790: gap of 100 bp 92791 122716: contig of 29926 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 868 967: gap of 100 bp 968 2133: contig of 1166 bp in length 2134 223: gap of 100 bp 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10880 10979; gap of 100 bp 10980 15646; contig of 4667 bp in length 15647 15746; gap of 100 bp 15747 25208; contig of 9462 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            867: contig of 867 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /clone_lib="RPCI-11 Human Male BAC"
1. .867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Center code: WIBR
Web site: http://www-seq.wi.mit.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1. .165147
/organism="Homo sapiens"
/db_xref="taxon:9606"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /note="assembly_fragment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /clone="RP11-2413"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          clone_end:SP6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                                                                                             COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Shreap, Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F., Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F., Boguslavkiy, L., Boukhgalter, B., Brown, A., Burkett, G., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J., Ferrelra, P., FitzHugh, M., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Graft, G., Hagos, B., Haeford, A., Horton, L., Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Landers, T., Lehoczky, J., Levine, R., Lieu, C., Liu, G., Locke, K., McPheeters, R., Medrim, J., Meneus, L., McCurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Morrow, J., Naylor, J., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AC021342 165147 bp DNA linear HTG 09-SEP-2000
Homo sapiens clone RPl1-2413, WORKING DRAFT SEQUENCE, 14 unordered
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1679 gcactggaaattgcttacaggaaatcctggctacagcgactgggtcccccctcagtctt 1738
                                                                                                                                                                                                             cagcaacctgggaccgaaccggtggcacaatgaacggatcacctcgtcaactggcaaaag 1798
                                                                                                                                                                                                                                                                                                                                                                                                           tgtatggcttttcaacatcacagccgacccatatgagagggtggacctatctaacaggta 1858
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9119 TGTATGGCTTTTCAACATCACAGCCGACCCATATGAGAGGGTGGACCTATCTAACAGGTA 9060
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gcaaaagaaaagc-aaaaaaaagaagaagaagaaacagcagaaagcagtctcaggttcaacttg 2097
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ccattcaggtgttacttgtgggataagcacaaatatttcctgtttggttaaactttaatca 2157
                                                          9239 GCACTGGAAATTGCTTACAGGAAATCCTGGCTACAGCGACTGGGTCCCCCCTCAGTCTTT
                                                                                                                                                                                                                                                             9179 CAGCAACCTGGGACCGAACCGGTGGCAAATGAACGGATCACCTTGTCAACTGGCAAAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          accatggtatagaggaaaccaagaaaaagaagccaagcaaaaatcaggctgagaaaa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 (bases 1 to 16514)
Birren, B., Linton, L., Nusbaum, C. and Lander, E. Homo sapiens, clone RP11-2413
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       gcctaagcgtcaggcttgttttcatgctgtgccac 2252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8699 GCCTAAGCGTCAGGCTTGTTTTCATGCTGTGCCAC 8665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HTG; HTGS_PHASE1; HTGS_DRAFT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AC021342
AC021342.3 GI:10045513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        pieces
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AC021342/c
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9059
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                   1739
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8879
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2098
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2218
                                                                                                                                                                                                                                                                                                                                                                                                           1799
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2039
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SOURCE
```

us-09-495-823-8

```
AC091322 232951 bp DNA linear HTG 11-JUN-2001
Mus musculus clone RP23-60D3, WORKING DRAFT SEQUENCE, 5 unordered
                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 5
AC091322/c
                                                                                                                                                                           1679
                                                                                                                                                                                                                                                                                                   1979
                                                                                                   1499
                                                                                                                           1559
                                                                                                                                                                                       100923
                                                                                                                                                                                                                                                                           1919
                          1319
                                                   1379
                                                              101223
                                                                                                              101103
                                                                                                                                                   1619
                                                                                                                                                                                                               100863
                                                                                                                                                                                                                           1799
                                                                                                                                                                                                                                       100803
                                                                                                                                                                                                                                                    1859
                                                                                                                                                                                                                                                                                                               100623
                                                                                                                                                                                                                                                                                                                                                                            2158
  1259
                                                                           1439
                                                                                                                                       101043
                                                                                                                                                               100983
                                                                                                                                                                                                   1739
                                                                                                                                                                                                                                                                                       100683
                                                                                                                                                                                                                                                                                                                           2039
                                                                                                                                                                                                                                                                                                                                        100563
                                                                                                                                                                                                                                                                                                                                                                                        100443
                                                                                                                                                                                                                                                                                                                                                                                                                100383
              g
                                      Ω
                                                              a
                                                                                      qq
                                                                                                              qq
                                                                                                                                      qq
                                                                                                                                                              g
                                                                                                                                                                                      qq
                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                               Dp
                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                       Dp
                                                                                                                                                                                                                                                                                                                                                                                                                g
                           δ
                                                   ò
                                                                           ŏ
                                                                                                   ò
                                                                                                                           à
                                                                                                                                                   δ
                                                                                                                                                                           ò
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                      ij
                                                                                                                                                                                                                                                                                                                                                                                                                                 101764
                                                                                                                                                                                                                                                                                                                     101644
                                                                                                                                                                                                                                                                                                                                                                                              101464
                                                                                                                                                                                                                                                                                                                                                          1078
                                                                                                                                                                                                                                                                                                                                                                                 1138
                                                                                                                                                                                                                                                                                                                                                                                                          1198
                                                                                                                                                                                                                                                                                                         928
                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                        atttgatacctttttggttcccttttgggaagtggggattactatacacactacaaatg
                                                                                                                                                                                                                      ä
                                                                                                                                                                                                          165147;
                                                                                                                                                                                 others
                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                          Length
                                                                                                                                                                                 1303
                                                                                                                                                                                                                     .;
S
                                                                                                                                                                                                         ;
;
                                                                                                                                                                                                          DB
                                                                                                                                                                                 ų
                                                                                                                                                                                                         Score 1515; DB
Pred. No. 0;
0; Mismatches
                                          /note="assembly_fragment"
92791. .122716
/note="assembly_fragment"
122817. .149492
/note="assembly_fragment"
149593. .165147
vector_side:left"
968. .2133
/note="assembly_fragment"
              2234. .3987
/note="assembly_fragment"
4088. .6971
                         4088. .6971
/note="assembly_fragment"
7072. .10879
                                                                                                                                                                                 51853
                                                                                                                                                               'note="assembly_fragment
                                                                                                                                                                          vector_side:right"
29805 c 30056 g
                                                                                                                                                                                                         66.9%;
99.6%;
                                                                                                                                                                      clone_end:T7
                                                                                                                                                                                                              Best Local Similarity 99.6
Matches 1529; Conservative
                                                                                                                                                                                 ಥ
                                                                                                                                                                                 52130
        misc_feature
                    misc_feature
                                            misc_feature
                                                       misc_feature
                                                                     misc_feature
                                                                                 misc_feature
                                                                                            misc_feature
                                                                                                         misc_feature
                                                                                                                    misc_feature
                                                                                                                                 misc_feature
                                                                                                                                            misc_feature
                                misc_feature
                                                                                                                                                                                                          Query Match
                                                                                                                                                                                 BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                 839
                                                                                                                                                                                                                                                                                             101763
                                                                                                                                                                                                                                                                                                         899
                                                                                                                                                                                                                                                                                                                    101703
                                                                                                                                                                                                                                                                                                                                 959
                                                                                                                                                                                                                                                                                                                                             101643
                                                                                                                                                                                                                                                                                                                                                                                 1079
                                                                                                                                                                                                                                                                                                                                                                                                          1139
                                                                                                                                                                                                                                                                                                                                                                                                                                  1199
                                                                                                                                                                                                                                                                                                                                                                                                                                             101403
                                                                                                                                                                                                                                                                                                                                                                                              101523
                                                                                                                                                                                                                                                                                                                                                                                                                      101463
                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                 à
```

```
100804
                                                                                101224
                                                                                                                                                                                                                                                                                                      100984
                                                                                                                                                                                                                                                                                                                                                                                                                  100864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100564
                                                                                                                                                                                                                                                                                                                                1678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1918
                                                                                                                                                                                                                       1558
                                                                                                                                                                                                                                                                                                                                                                                        1738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1978
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gacttatggtttctataacaacagcattatcatttactcttcagataatggtggccagcc
                                                                                                           101163 GGAACTTGTGCACATCACTGGTACCCCACTCTCATTTCACTGGCTGAAGGACAGAT
                                                                                                                                                                                                                                                                           gcactggaaattgcttacaggaaatcctggctacagcgactgggtcccccctcagtcttt
                                                                                                                                                                                                                                                                                                                                                                                                     100743 TCCAGGAATCGTGAAGAAGCTCCTACGAGGCTCTCACAGGTTCAACAAACTGCAGTGCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GCAAAAGAAAAGAAAAAAAAAAAGAAGAAACAGCAGAAAGCAGTCTCAGGTTCAACTTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           101343 GACTTATGGTTTCTATAACAACAGCATTATCATTTACTCTTCAGATAATGGTGGCCAGCC
                                                                                                                                                                                                                                     101283 TACGGCAGGAGGAGTAACTGGCCTCTCAGAGGTAGCAAAGGAACATATTGGGAAGGAGG
                                                                                                                                                                  ggaacctgtgcacatcactgactggtaccccactctcatttcactggctgaaggacagat
                                                                                                                                                                                                                       tgatgaggacattcaactagatggctatgatatctgggagaccataagtgagggtcttcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        tecaggaategtgaagaageteetaeggaggeteteaeagtteaaeaaaetgeagtgee
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       gcaaaagaaaagc-aaaaaaagaagaagaaacagcagaaagcagtctcaggttcaacttg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gttettatettteatetgttteetaggtaaaceageaaatttggetegataatategetg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GCCTAAGCGTCAGGCTTGTTTCATGCTGTGCCAC 100349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2218 gcctaagcgtcaggcttgtttcatgctgtgccac 2252
```

~

```
ų
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No. 0;
0; Mismatches
                                                                                                                                                                                                                                                      /note-"assembly_fragment"
2691. 11774
/note-"assembly_fragment"
11875. 110624
/note-"assembly_fragment"
100625. 232951
/note-"assembly_fragment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   50.1%; Score 1134.2; 84.7%; Pred. No. 0;
                                                                                                                                                                                                                                                                                                                                                                                           vector_side:right"
45474 c 46842 g 69434
                                                                                                                                                                                           /note="assembly_fragment
                                                                                                                                                                                                                        vector_side:left"
154. .2590
                                                                                                                                                                                                           clone_end:SP6
                                                                                                                                                                                                                                                                                                                                                                              clone_end:T7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 84.7
Matches 1303; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                  ď
                                                                                                                                                                                                                                                                                                                                                                                                                  70795
                                                                                                                                                                           misc_feature
                                                                                                                                                                                                                                             misc_feature
                                                                                                                                                                                                                                                                               misc_feature
                                                                                                                                                                                                                                                                                                             misc_feature
                                                                                                                                                                                                                                                                                                                                              misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                          source
                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
                                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ολ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ...... Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as some as it is available and the accession number will be preserved.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Insert size: 210000; agarose-fp
Insert size: 23251; sum-of-contigs
Quality coverage: 10.7 in Q20 bases; agarose-fp
Quality coverage: 9.6 in Q20 bases; sum-of-contigs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequencing vector: Plasmid; n/a; 100% of reads Chemistry: Dye-terminator Big Dye; 100% of reads Assembly program: Phrap; version 0.960731 Consensus quality: 231176 bases at least Q40 Consensus quality: 231965 bases at least Q30 Consensus quality: 232319 bases at least Q20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 53: contig of 53 bp in length 54 153: gap of 100 bp 154 2590: contig of 2437 bp in length 2591 2690: gap of 100 bp
                                                                                                                                   Birren, B., Linton, L., Nusbaum, C. and Lander, E. Mus musculus, clone RP23-60D3
ÀCO91322.
ACO91322.2 GI:14336533
HTG; HTGS_PHASE1; HTGS_FULLTOP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Center project name: L132
Center clone name: 60_D_3
                                                                                                                        (bases 1 to 232951)
                                                                     Mus musculus
                                                     house mouse.
                                                                                                                                                                          Unpublished
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                   ORGANISM
                                                                                                                   REFERENCE
AUTHORS
TITLE
                                                                                                                                                                      JOURNAL
REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE
JOURNAL
```

```
Db 147754 TTCCTCAAAGGTATCAGATACACACGGGACTTCAGCATTCTATCATTAGACCTACCCAAC 147695
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Db 147574 AGAGAGGATTTGATACCTTTTTTGGCTCTCTTTTGGGAAGTGGTGATTATTATACACACT 147515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DD 147454 GGGACTATGACAATGGCATCTATTCAACACAGATGTACACGCAGAGAGTGCAGAGAAATCT 147395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DD 147394 TAGCCACCCATGATCCCACAAAACCTCTATTTTTATATGTGGCTTACCAAGCTGTCCACT 147335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Db 147334 CACCACTTCAAGCCCCCTGGCAGGTACTTTGAACACTACAGATCCATTATCAACATAAACA 147275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DD 147214 CCCTAAAGAGGTATGGTTTCTATAACAATAGCATTATCATTTACTCCTCAGATAATGGTG 147155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1072 tagetteccataaccecacaaageetatattttatatattgeetateaagetgtteatt 1131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1132 caccactgcaagctcctggcaggtatttcgaacactaccgatccattatcaacataaaca 1191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Qy 1192 ggaggagatatgctgccatgctttcctgcttagatgaagcaatcaacaacgtgacattgg 1251
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1252 ctctaaagacttatggtttctataacaacagcattatcatttactcttcagataatggtg 1311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1012 gggactatgacaatggcatatactccacacagatgtacactcagagagtacagcaaatct 1071
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      712 ttactggaaagtatcagatacacaccggacttcaacattctatcataagacctacccaac 771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Db 147274 GGAGGAGGTACGCACCCATGCTTCTGCATGAAGCAATCCACAATGTGACCCTGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 232951;
2691 11774: contry v. 100 bp
11775 11874: gap of 100 bp
11875 10024: contig of 88650 bp in length
100525 100624: app of 100 bp
100625 232951: contig of 132327 bp in length.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    406 others
                                                                                                                                                     DB 2;
```

ä

8

```
ctcagtctttcagcaacctgggaccgaaccggtggcacaatgaacggatcacctcgtcaa 1788
 Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6
                                                                                                                                                                                                                                                                                                                                      /clone="LNGOB487"
/tissue_type="human lung"
/clone_lib="LNG"
/note="cloning vector pMEIBSFL3"
204. .713
                                                                                                                                                                                                                                                                                                                                                                                            /note="unnamed protein product"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 789.8; DB
Pred. No. 2e-226;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 546
                                                                                                                                                                                                                                                                                                                 /organism="Homo sapiens"
/db_xref="taxon:9606"
                                                                                                                                                                                                                                                                                            Location/Qualifiers
1. .1871
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 354
                                                                                                                                                                                                                                                                                                                                                                                                      /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          34.9%;
98.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TCHSGVTCG"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                367 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches 816; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ø
                                                                                                                                                                                                                                                                                     Agency).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BASE COUNT
ORIGIN
  ORGANISM
                               REFERENCE
AUTHORS
                                                                                               TITLE
JOURNAL
REFERENCE
                                                                                                                                                    TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1729
                                                                                                                               AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                   CDS
                                                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                    COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRI 29-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          oligo capping; fis (full insert sequence).
Homo sapiens human lung cDNA to mRNA, clone_lib:LNG clone:LNGO8487.
                                                                                                                                                                       147095
                                                                                                                                          CCGTCCCTGTCAGGTACCCCCGAAGGATCCCAGAAGCAACCCTAGGCTCAACGGAGGAG 146495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1431
                                                                                    1491
                                                                                                                                                                                                                                                                                                      1791
                                                                                                                                                                                                                                                                                                                                                  1851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2147
                                                                                                                              gacagattgatgaggacattcaactagatggctatgatatctgggagaccataagtgagg 1551
                                                                                                                                                                                                                                                                                                                                                                                            1911
                                                                                                                                                                                                                                                                                                                                                                                                                                      cagigccggicaggiatccccccaaagaccccagaagiaaccctaggcicaaiggagggg 1971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2148 actttaatcagttcttatctttcatctgtttcctaggtaaaccagcaaatttggctcgat 2207
                                                                                                                                                                                                                                                                                                                                                                                           147154 GGCAGCCCACAGCAGGAAGGAAGTAACTGGCCACTCAGAGGCAGCAAAGGAACATATTGGG
                                                     tgtgtaaggaacctgtgcacatcactgactggtaccccactctcatttcactggctgaag
                                                                                                                                                                                                                                                                                                                                                 gcaaaagtgtatggcttttcaacatcacagccgacccatatgagagggtggacctatcta
 gccagcctacggcaggagggagtaactggcctctcagaggtagcaaaggaacatattggg
                                           aaggaaggatccgggctgtaggctttgtgcatagcccacttctgaaaaacaagggaacag
                                                                                                                                                                                                                                                                                                       agtettteageaacetgggaeeggaeeggtggeaeaatgaaeggateaeetegteaaetg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        agaaaaagcaaaagcaaaa----aaagaagaagaaacagcagaaagcagtctcag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gttcaacttgccattcaggtgttacttgtggataagcacaaatatttcctgtttggttaa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1871 bp mRNA linear
Homo sapiens CDNA: FLJ23548 fis, clone LNG08487.
AK027201.
Gligo AK077201. GI:10440271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                            1672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2088
1312
                                         1372
                                                                                    1432
                                                                                                                             1492
                                                                                                                                                                        1552
                                                                                                                                                                                           146914
                                                                                                                                                                                                                                                                                                       1732
                                                                                                                                                                                                                                                                                                                                                 1792
                                                                                                                                                                                                                                                                                                                                                                                                                                     1912
                                                                                                                                                                                                                                                                                                                                                                                                                                                         146554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2032
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Db 146268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AK027201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LOCUS
                    g
                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                        g
                                                                                                                                                g
                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
ò
                                         à
                                                              g
                                                                                   õ
                                                                                                                             ð
                                                                                                                                                                       ð
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                      ô
                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

/protein\_id="Bab15689.1"
/db\_xref="G1:10440272"
/db\_xref="G1:10440272"
/db\_xref="G1:10440272"
/db\_xref="G1:10440272"
/db\_xref="G1:10440272"
/db\_xref="G1:10440272"
/db\_xref="G1:10440272"
/db\_xref="G1:10440272"
/db\_xref="G1:1040272"
/db\_xref="G Human Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T., Shibahara, T., Tanaka, T. and Nakamura, Y.
Direct Submission
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:chaleims.u-tokyo.ac.jp, Tel:81-3-5449-5286, 7 1 (sites)
Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,
Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Tanigami,A.,
Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M.,
Ohmori,Y., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T.,
Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNa sequencing project
Unpublished (2000)
2 (bases 1 to 1871) Euteleostom1; NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan: cDNA full insert sequencing; Research Association for Blotechnology: cDNA library construction, 5'- & 3'- and one pass sequencing: Departent of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology 1728 1490 aggacagattgatgaggacattcaactagatggctatgatatctgggagaccataagtga 1549 1609 1610 aaaaaatggctcctgggcagcaggctatgggatctggaacactgcaatccagtcagccat 1669 138 Gaps 19 AGTGTGTAAGGAACTTGTGCACATCACTGACTGGTACCCCACTCTCATTTCACTGGCTGA 78 1430 agtgtgtaaggaacctgtgcacatcactgactggtaccccactctcatttcactggctga 199 AAAAAATGGCTCCTGGGCAGGCTATGGGATCTGGAACACTGCAATCCAGTCAGCCAT 5 Length 1871;

ò a ò 엄 ò g ö a

```
59627
59727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckerly, R., Boguslavkiy, L., Boukhgalter, B., Brown, A., Castle, A., Collangelo, M., Collymore, A., Cooke, P., Castlano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, M., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Mardias, A., Klein, J., Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J., Marlor, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HTG 07-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 157043)

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

Homo sapiens chromosome, clone RP11-21L8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Submitted (13-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Dec 7, 2000 this sequence version replaced gi:10280848.
All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACU13692 157043 bp DNA linear HTG 07-DEC-2C HOmo sapiens clone RPl1-21L8, *** SEQUENCING IN PROGRESS ***, 45 unordered pieces.
                                                                                                                                                                                                                                                                                                                                                              gggtctagggaccatggtatagagagagaaaccaagaaaaagaagccaagccaaaccagcaaatcagg 2028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           gttcaacttgccattcaggtgttacttgtggataagcacaaatatttcctgtttggttaa 2147
                                                                                                                                                                   ctaacaggtatccaggaatcgtgaagaagctcctacggaggctctcacagttcaacaaaa 1908
                                                                                                                                                                                                                                                              ctgcagtgccggtcaggtatcccccaaagaccccagaagtaaccctaggctcaatggag 1968
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ctgagaaaaagcaaaagaaaagc-aaaaaaagaagaagaaacagcagaaagcagtctcag 2087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GTTCAACTTGCCATTCAGGTGTTACTTGTGGATAAGCACAAATATTTCCTGTTTGGTTAA 738
                                                                                                                                                                                          ctggcaaaagtgtatggcttttcaacatcacagccgacccatatgagagggtggacctat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AC013692.5 GI:11597077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (bases 1 to 157043)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HTG; HTGS_PHASE1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AC013692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 7
AC013692/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
                                                                                                                                                              1849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        739
                                                                    1789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM
                    319
                                                                                                                                                                                                                                                              1909
                                                                                                                                                                                                                                                                                                                                                           1969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TITLE
JOURNAL
REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2088
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LOCUS
```

음

q ò g ò

ò

g

ò

```
......... Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
                                                                                                                                         NOTE: This is a 'working draft' sequence. It currently consists of 45 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
http://ftp.genome.washington.edu/RM/RepeatMasker.html
                                    8631: contig of 1599 bp in length
8791: gap of 100 bp
10237: contig of 1446 bp in length
10337: gap of 100 bp
12115: contig of 1778 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  p of 100 bp contig of 2793 bp in length p of 100 bp contig of 2088 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          o of 100 bp
contig of 2806 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           p of 100 bp contig of 3213 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26: gap of 100 bp
62247: contig of 2521 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     oof 100 bp contig of 1659 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  p of 100 bp
contig of 1401 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             oof 100 bp
contig of 1865 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              o of 100 bp contig of 3122 bp in length
                                                                                                                                                                                                                                                                           925 1024; gap of 100 bp 1025 1800; contig of 776 bp in length 1801 1900; gap of 100 bp in length 1901 2800; contig of 900 bp in length
                                                                                                                                                                                                                                                                                                       100 bp
f 900 bp in length
                                                                                                                                                                                                                                                               924: contig of 924 bp in length
                                                                                                                                                                                                                                                                                                                                                                                   854 bp in length
                                                                                                                                                                                                                                                                                                                                                      728 bp in
                                                                                                                                                                                                                                                                                                                                                                                                                            100 bp
of 1129 bp i
100 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98: gap of 100 bp
25224: contig of 2326 bp
24: gap of 100 bp
27369: contig of 2045 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              oof 100 bp
contig of 2265 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         of 100 bp
contig of 1620 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      16201 16300: gap of 100 bp
16301 17801: contig of 1501 bp
17802 17901: gap of 100 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                76: gap of 100 bp 50235: contig of 9659 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       56870: gap of 100 bp 59626: contig of 2756 bp
                                                                                                                                                                                                                                                                                                                                                                                               of 100 bp
contig of 1081 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61: gap of 100 bp
22798: contig of 1637 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    oof 100 bp
contig of 3055 bp
                                                                                                                                                                                                                                                                                                                                     100 bp
                                                                                                                                                                                                                                                                                                                                                                    dq 001
                                                                                                                                                                                                                                                                                                                                                     contig of
                                                                                                                                                                                                                                                                                                                                                                                contig of
                                                                                                                                                                                                                                                                                                                                                                                                                                          contig of
                                                                                                                 Center clone name: 21_L_8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             362: gap of 32450: cont
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10338 12115: cont
12116 12215: gap of
12216 14480: cont
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      14481 14580: gap of
14581 16200: cont
16201 16300: qap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    37421: gap of
40476: cont
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  27469: gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          32550: gap of
34415: con
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       34515: gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          53557: gap of
                                                                                                                                                                                                                                                                                                                                                                                                                           3: gap of
6992: con
2: gap of
8691: con
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 40576: gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    59726: gap of
                                                                                                                                                                                                                                                                                                                                     2900: gap of
3628: cor
                                                                                                                                                                                                                                                                                                                                                                    3728: gap of
                                                                                                                                                                                                                                                                                                                                                                                               4682: gap of
5763: co
                                                                                                                                                                                                                                                                                                                                                                                   4582:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21061:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        37321:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19560:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           56770:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 30262:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              53457:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22898:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19660:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25324:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              50335:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30362:
                                                                                                                                                                                                                                                                                                                                                                                                                              5863:
                                                                                                                                                                                                                                                                                                                                                                                                                                                          7092:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21161
                                                                                                                                                                                                                                                                                                                                     2801
3629
3729
4583
4683
5764
5864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 17902
19561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19661
21062
21162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22799
22899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 40577
50236
50336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          53458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7093
8692
8792
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               27470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 40477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           53558
56771
56871
```

```
69112 59114: gap of 100 bp 69115 13176: contig of 4062 bp in length 73177: 73276: gap of 100 bp 73177: 73276: gap of 100 bp 73277 73276: gap of 100 bp 78379 79079: gap of 100 bp 8339 83497: gap of 100 bp 8798 87887: contig of 4339 bp in length 8798 87887: contig of 4339 bp in length 8798 87987: gap of 100 bp 8798 91530: gap of 100 bp 91531 91530: contig of 5834 bp in length 103022 103121: gap of 100 bp 100 bp 113421 113420: contig of 5835 bp in length 113421 113420: gap of 100 bp 113421 113420: gap of 100 bp 113421 113420: gap of 100 bp 1130195 130294: gap of 100 bp 114726: contig of 9169 bp in length 137957 147125: contig of 9169 bp in length 137957 147125: contig of 7741 bp in length 154926 155066: gap of 100 bp 14726 155066: gap of 100 bp 14726 155066: gap of 100 bp 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 155067 1550
3362 bp in length
                        p of 100 bp
contig of 3205 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                clone_lib="RPCI-11 Human Male BAC"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12216. .14480
/note="assembly_fragment"
14581. .15200
/note="assembly_fragment"
16301. .17801
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1025. 1800
/note="assembly_fragment"
1901. 2800
/note="assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         093. .8691
note="assembly_fragment"
792. .10237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 s/92. .10237
'note="assembly_fragment"
0338. .12115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2901. 3628
/note="assembly_fragment"
3729. 4582
/note="assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note="assembly_fragment"
7093. .8691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0338. 12115
note="assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /note="assembly_fragment"
19661. .21061
/note="assembly_fragment"
21162. .22798
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         683. .5763
note="assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      note="assembly_fragment"
7902. .19560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note="assembly_fragment
clone_end:SP6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1. .157043
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-21L8"
                                                                             69114: gap of 1
73176: contig of
65709: contig of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ector_side:left"
         ... gap of
69014:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .6992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURES
```

```
ï

      Qy
      82
      ggaggaggaggagaaaqtgaaatgtgctggagaaggggagcctccttgttcttccgga
      141

      bb
      110231
      GGAGGAGGAGGAGAAAGTGAAATGTGCTGGAGAAAGGCGAGCCCTCCTTGTTCTTCCGGA
      110172

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Db 110290 GTGGAGATATTAACTTTTTTC-TTTTTTTTTTTGGTGGAAGCTGCTCTAGGGAGGGG 110232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Db 109811 CTCAAGCTGGAGAAACTAGAGCCCAGCACAACTTCCACCTCCCAGCCCCATCTCATT 109752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Db 109931 GTCCTGGAAAGATGCTAGCAATGGGGGCGCTGGCAGGATTCTGGATCCTCTCTGCTCTCA 109872
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Db 109871 CTTATGGTTACCTGTCCTGGGGCCAGGCCTTAGAAGAGGAGGAAGAAGGGGCCCTTACTAG 109812
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    681
                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  202 agaggaggaggatttcttgccaggtggagagtcttcaccgtctgttgggtgcatgtgt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          262 gegecegeaseggegeggggegegtggtteteegegtggagteteaectgggaeetgagt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             502 ctcaagctggagagaaactagagcccagcacaacttccactcccagcccatctcattt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      cacctactcttgacaagctcgctgccgaaggagttaaactggagaactactatgtccagc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      382 gtcctggaaagatgctagcaatgggggcgctggcaggattctggatcctctgcctcctca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              cttatggttacctgtcctggggccaggccttagaagaggaggaagaaggggccttactag
                                                                                                                                                                                                                                                                                                                                                                                                                            Length 157043
                                                                                                                                                                                                                                                                                                                                                                                          1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                   Query Match 30.5%; Score 690.6; DB 2; Best Local Similarity 99.0%; Pred. No. 5.6e-196; Matches 704; Conservative 1; Mismatches 5;
                                                                                                                                             32551. 34415 - 7016="assembly_fragment" 34516. 37321 /note="assembly_fragment" 37422. 40476 /note="assembly_fragment" 40577. 50235 /note="assembly_fragment"
                                                                      27470. 30262 /note="assembly_fragment" 30363. 32450 /note="assembly_fragment"
22899. .25224 '-
/note="assembly_fragment"
25325. .27369
                                                   /note="assembly_fragment"
                                                                                                                                                                                                                                                                                            50336. .53457
/note="assembly_fragment"
   misc_feature
                                      misc_feature
                                                                      misc_feature
                                                                                                                                                   misc_feature
                                                                                                                                                                                                                       misc_feature
                                                                                                                                                                                                                                                            misc_feature
                                                                                                                                                                                                                                                                                                misc_feature
                                                                                                              misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δλ
                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δλ
```

DEFINITION

AC021342

ACCESSION VERSION KEYWORDS SOURCE

ORGANISM

REFERENCE AUTHORS

TITLE

JOURNAL REFERENCE AUTHORS

```
46718 60565: contig of 13848 bp in length 60566 60665: gap of 100 bp 60666 92690: contig of 32025 bp in length 92691 92790: gap of 100 bp 92791 122716: contig of 29256 bp in length 12277 122816: gap of 100 bp 122817 149492: contig of 26676 bp in length 149493 149592: gap of 100 bp 149592: gap of 100 bp 149593: gap of 100 bp 149593: gap of 100 bp 149593: gap of 100 bp 169592: gap of 100 bp 169592: gap of 100 bp
                                                                                                                                                                                                         1303 others
                                                                                                                               10880 10979: gap of 100 bp
10980 15646: contig of 4667 bp in length
15647 15746: gap of 100 bp
15747 25208: contig of 9462 bp in length
                                                                                 of 100 bp
contig of 3808 bp in length
p of 100 bp
contig of 4667 bp in length
     contig of 1754 bp in length of 100 bp contig of 2884 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1. .165147
/organism="Homo sapiens"
/db_xrsf="texnon:9606"
/clone="RP11-2413"
/clone=lib="RPCI-11 Human Male BAC"
1. .867
/note="assembly_fragment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       30.5%; Score 690.6; DB 2;
99.0%; Pred. No. 5.7e-196;
ive 1; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          vector_side:right"
29805 c 30056 g 51853 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           33559. .46617

/note="assembly_fragment"

46718. .60565

/note="assembly_fragment"

6066. .92690

/note="assembly_fragment"

92791. .122716

/note="assembly_fragment"

/note="assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2234. .3987
/note="assembly_fragment"
4088. .6971
/note="assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7072. .10879
/note="assembly_fragment"
10980. .15646
/note="assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    'note="assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15747, .25208 -- 'note="assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     'note="assembly_fragment"
3 4087: gap of 100
3 4087: gap of 100
3 671: contig of 21
5 7071: gap of 100
2 10879: contig of 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           clone_end:T7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 704; Conserv
        2234
3988
4088
6972
7072
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     52130 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                             Uppublished

State (Dases 1 to 165147)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Bedde, F., Bodukhgalter, B., Brown, A., Burkett, G., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J., Ferrelra, P., FitzHugh, M., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Macdonald, P., Marquis, M., Mores, R., Liud, G., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McMenta, C., Liud, G., Locke, K., McPheeters, R., Meldrim, J., Meneus, L., McGurk, A., McKernan, K., Pierre, N., Pisanl, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

In separat, Submission

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Sep 9, 2000 this sequence version replaced gi:6899755.

All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1997)

http://ftp.genome.washington.edu/RW/RepeatMasker.html
                                                           HTG 09-SEP-2000
                                                      ACO21342 165147 bp DNA linear HTG 09-SEP-2001
Homo sapiens clone RP11-2413, WORKING DRAFT SEQUENCE, 14 unordered
                                                                                                                                                                                                                                                             Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------ Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NOTE: This is a 'working draft' sequence. It currently consists of 14 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Center project name: 14.177
Center clone name: 24.1.3
Center clone name: 24.1.3
Center clone name: 24.1.3
Sequencing vector: M13; M77815; 100% of reads Sequencing vector: M13; M77815; 100% of reads Assembly program: Phrap; version 0.960731
Consensus quality: 156719 bases at least 040
Consensus quality: 160377 bases at least 030
Consensus quality: 162195 bases at least 030
Insert size: 170000; agarose-fp
Insert size: 170000; agarose-fp
Insert size: 170000; agarose-fp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Quality coverage: 4.4 in Q20 bases; agarose-fp Quality coverage: 4.6 in Q20 bases; sum-of-contigs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
------ Project Information
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 867: contig of 867 bp in length
868 967: gap of 100 bp
968 2133: contig of 1166 bp in length
2134 2233: gap of 100 bp
                                                                                                                                                                                                                                                                                                1 (bases 1 to 165147)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-2413
                                                                                                                                                             AC021342.3 GI:10045513
HTG; HTGS_PHASE1; HTGS_DRAFT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Center code: WIBR
                                                                                                                                                                                                                                            Homo sapiens
                                                                                                          pieces.
AC021342
AC021342.3
                                                                                                                                                                                                                 human.
```

TITLE JOURNAL

COMMENT

1;

Indels

Length 165147;

Wed

ò පි 셤 ò g

õ

QQ

ò

g ò g ô g ò ద ò g ö g à

ò

```
Contact: submissions@watson.wustl.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                 /clone="RP11-219L11"
                 ------- Project
                                                                                                                                                                                                                                                                                                                                                                                                                                      /chromosome="4"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity 99.0 nes 704; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 54056 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         91008
                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AC105414 175645 bp DNA linear HTG 09-FEB-2002 Homo sapiens chromosome 4 clone RP11-219L11, WORKING DRAFT SEQUENCE, 2 unordered pieces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Submitted (04-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
On Feb 9, 2002 this sequence version replaced g1:18483567.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 175645) Waterston, R.H.
                                                                                                                                                                                                                                    40066
                                                                                                                                                                                                                                                                39886
                                                                                                                                39946
                                                                                                                                                                                                                                                                                                                              40246
                                                                                                                                                                                                                                                                                                                                                                                                                                  40006
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   40366
                                                                                                                                                                                                                                                                                                                                                                    561
                                                    201
                                                                                                        261
                                                                                                                                                          321
                                                                                                                                                                                                             381
                                                                                                                                                                                                                                                                                                                  501
                                                                                                                                                                                                                                                                                                                                                                                                                       621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         681
  gaatggctcccaggggctgtgcgggcatccgccttctccacaggcctgtgtct
                                                                                                                                                                                                                                    40007 GAATGGCTCCCAGGGGCTGTGCGGGCATCCGCCTTCTCCCACAGGCCTGTGTCT
                                                                                                                                                                                                                                                                                                                                                                    ctcaagctggagagaaactagagcccagcacaacttccacctcccagccccatctcattt
                                                                                                                                                                                                                                                                                                                                                                                                                       tcatcctagcggatgatcagggatttagagatgtgggttaccacggatctgagattaaaa
                        39767 GGAGGAGGAGGAGAAAGTGAAATGTGCTGGAGAAGAGGGAGCCCTCCTTGTTCTTCCGGA
                                                   gtcccatccattaagccatcacttctggaagattaaagttgtcggacatggtgacagctg
                                                                                                                                                                                                                                                                                                                  63108, USA
Feb 9, 2002 this sequence version replaced gi:18483567
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACI05414
AC105414.4 GI:18642912
HTG; HTGS_PHASE1; HTGS_FULLTOP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The sequence of Homo sapiens clone Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Center
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 (bases 1 to 175645)
Waterston, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 9
AC105414/C
LOCUS
DEFINITION
                                                   142
                                                                                                                                                                                                                                                                                                                                                                                                                       562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SOURCE
ORGANISM
82
                                                                                                                                                                                                                                                              382
                                                                                                                                                                                                                                                                                                                442
                                                                                                      202
                                                                                                                               39887
                                                                                                                                                                                                           322
                                                                                                                                                                                                                                                                                                                                          40127
                                                                                                                                                                                                                                                                                                                                                                    502
                                                                                                                                                                                                                                                                                                                                                                                            40187
                                                                                                                                                                                                                                                                                                                                                                                                                                              40247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMMENT
```

```
91249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              91129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                * NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gegecegeaseggegegggggegtggtteteegegtggagteteacetgggaeetgagt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   91068 GCGCCCGCAGCGCGCGCGCGCGTGCTTCTCCGCGTGGAGTCTCACCTGGGACCTGAGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gaatggctcccaggggctgtgcggggcatccgcctccgccttctccacaggcctgtgtct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gtcctggaaagatgctagcaatgggggcgctggcaggattctggatcctctgctcctca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 175645;
                                                                                                     Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 174714 bases at least 040
Consensus quality: 174877 bases at least 020
Insert size: 177000; agarose-fp
Insert size: 175545; sum-of-contigs
Quality coverage: 10.91 in Q20 bases; sum-of-contigs
Quality coverage: 10.91 in Q20 bases; sum-of-contigs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 74211: contig of 74211 bp in length 74212 74312: gap of unknown length 74312 175645: contig of 101334 bp in length Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30.5%; Score 690.6; DB 2; 99.0%; Pred. No. 5.8e-196; iive 1; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    74312. .175645 / Anote="assembly_name:Contig35" 32068 c 32415 g 56997 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  'note-"assembly_name:Contig34"
                                        Sequencing vector: M13; 0%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1.175645
/organism="Homo sapiens"
/db_xref="taxon:9606"
Center project name: H_NH0219L11
                                                                                      Sequencing vector: plasmid; 100%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

1;

989

969

756 619 816 739 876 799 936 859

```
attactatacacactacaaatgtgacagtcctgggatgtgtgggctatgacttgtatgaaa 996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TGGGTCTGGCAGGTGCTACCACCTCAGCAGCCGATGGCTAGATGGCTACGACGTGTGGC
                                                                                                                                                                                  CGCGGAGCCAGCTCCTCACTGGCAGGTACCAGATCCACAGGACTCCAGCATTCCATCA
                                                                                                                                                                                                                    tttcactggctgaaggacagattgatgaggacattcaactagatggctatgatatctggg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          agaccataagtgagggtcttcgctcacccgagtagatattttgcataacattgaccca
                                     GA--GCAGCCCTCGGCCGCTCCCCCAGCTCCCCCACATCTTCATCCTCAGGGACG
                                                                        agctcgctgccgaaggagttaaactggagaactactatgtccagcctatttgcacacct
                                                                                                                                                                                                                                                                                  acgacaatgctgcctgggactatgacaatggcatatactccacacagatgtacactcaga
                                                                                                                                                                                                                                                                                                                                                                                                                              ttatcaacataaacaggaggagatatgctgccatgctttcctgcttagatgaagcaatca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----AAGCCGAGCTTCGTGGCCGACGGGCCCGGGGAGGCTGGC
                          aactagagcccagcacaacttccacctcccagccccatctcatttcatcctagcggatg
                                                                                                                                   GCTGGCGGCCAAGGGGGTCAAGTTGGAGAATTATTACATCCAGCCCATCTGCACGCCTT
                                                                                                                                                                                                                                                                      aggaggttggatattcaacgcatatggtcggaaaatggcacttgggtttttacagaaaag
                                                                                                                                                                                                                                                                                                                      aatgcatgcccaccagaagaggatttgataccttttttggttcccttttgggaagtgggg
                                                                                                                                                                                                                                                                                                                                                                                          ACTATTACACCTATGACAACTGTGATGGCCCAGGCGTGTGCGGCTTCGACCTGCACGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1117 atcaagctgttcattcaccactgcaagctcctggcaggtatttcgaacactaccgatcca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TCCAGGCAGTACACACACCCTGCAGTCCCCTCGTGAGTACCTGTACCGCTACCGCACCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1237 acaacgtgacattggctctaaagacttatggtttctataacaacagcattatcatttact
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               aaggaacatattgggaaggagggatccgggctgtaggctttgtgcatagcccacttctga
                                                                                                                                                                       ccaggagtcagtttattactggaaagtatcagatacacccggacttcaacattctatca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gagtacagcaaatcttagcttcccataaccccacaaagcctatattttatatattgcct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               cttcagataatggtggccagcctacggcaggagggagtaactggcctctcagaggtagca
  CCTGGGACTGGGCC-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1037
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                1057
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1277
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1397
                         517
                                                 442
                                                                                              200
                                                                                                                       637
                                                                                                                                             260
                                                                                                                                                                     697
                                                                                                                                                                                             620
                                                                                                                                                                                                                   757
                                                                                                                                                                                                                                            980
                                                                                                                                                                                                                                                                   817
                                                                                                                                                                                                                                                                                            740
                                                                                                                                                                                                                                                                                                                    877
                                                                                                                                                                                                                                                                                                                                           800
                                                                                                                                                                                                                                                                                                                                                                  937
                                                                                                                                                                                                                                                                                                                                                                                          860
                                                                                                                                                                                                                                                                                                                                                                                                                 997
                                                                                                                                                                                                                                                                                                                                                                                                                                        920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1097
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1537
                                            g
                                                                                        g
                                                                                                                                             a
                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                 P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                 δ
                                                                οy
                                                                                                                     δŽ
                                                                                                                                                                     Ω
                                                                                                                                                                                                                    δý
                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                   ρý
                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ILASHSPQRPLFLYVAFQAVHTPLQSPREYLYRTRTMGNVARRKYAAMVTCMDEAVRN
THAALKRYBEYNNSVIIFSSDNGGOFFSGSSMRPLGRKGTYWEGGYRGLGFVHSPLL
KTRGRTSRALMHITDWYPTLVGLAGGTTSAADGLOGYDWPALSEGRASPRTEILHNI
DPLYNHAQHGSLEGGFGIWNTAVQAAIRVGEWKLLIGDPGYGDWIPPQYLATFPGSWM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NLERWASVRQAVWLFNISADPYEREDLAGQRPDVVRTLLARLAEVNRTAIPVRYPAEN
PRAHPDFNGGAWGPWASDEEEEEEGRARSFSRGRRKKKCKICKLRSFFRKLNTRLMS
                                                                                                                                                                                                                                                                                                        PAT 30-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                           Euteleostomi;
GTCCTGGAAAGATGCTAGCAATGGGGGGCGCTGGCATTCTGGATCCTCTGCCTCCTCA 90889
                                                                                             90769
                                                                                                                                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostom Mammalai; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (Dases 1 to 2940) Glucksmann, M.A., Williamson, M., Rudolph-Owen, L.A. and Tsai, F.Y. Human sulfateses Patent: WO 0155411-A 6 02-AUG-2001; Millennium Pharmaceuticals, Inc. (US)
                                                                                                                                                                     681
                                                                                                                     621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   tagcaatggggggggctgggattctggatcctctgcctcctcacttatggttacctgt 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cctggggccaggccttagaagaagaagaagaagaaggagccttactagctcaagctggagaga 516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                      ctcaagctggagagaactagagcccagcacacttccacctcccagccccatct
                                                                                 cacctactcttgacaagctcgctgccgaaggagttaaactggagaactactatgtccagc
                                                                                                                     tcatcctagcggatgatcagggatttagagatgtgggttaccacggatctgagattaaaa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15;
                                                                                                                                                                                                                  Length 2940;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     _
_
_
_
                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 596.2; DB 6;
Pred. No. 5.7e-168;
0; Mismatches 598;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     product"
                                                                                                                                                                                                                                                                                                         DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /organism="Homo sapiens"
/db_xref="taxon:9606"
334. .2043
                                                                                                                                                                                                                                                                                                   Sequence 6 from Patent WO0155411.
AX206965
AX206965.1 GI:15394722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       334. .2043
/note="unnamed protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
1. .2940
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         б
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        902
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       26.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       915 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                             Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        567
                                                                                                                                                                                                                                                                                                                                                                  human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Sim
nes 1010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                             RESULT 10
AX206965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                               DEFINITION
                                                                                                                                                                                                                                                                                                                            ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                             ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                           AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Loc
Matches
                                                                                                                                                                   622
                                                                                                                                                                                                                   682
90948
                        442
                                                                     502
                                                                                                                     562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              329
                                             90888
                                                                                            90828
                                                                                                                                           90768
                                                                                                                                                                                           90708
                                                                                                                                                                                                                                          90648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      457
                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                  JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                       LOCUS
                                             q
                                                                                         g
                                                                                                                                                                                        q
                                                                                                                                                                                                                                        g
                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                        ò
                                                                     ò
                                                                                                                    ò
                                                                                                                                                                   ò
                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

1056

919

1116

1176

1276

```
412 (
                                                                                                                                                                                                                                                 942
                                                                                                                                                                                                                                                                     532
                                                                                                                                                                                                                                                                                                                                                                                              709 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                         1242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1422
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1009
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1189
                                                                                 702
                                                                                                      292
                                                                                                                                             352
                                                                                                                                                                                                                                                                                          1002
                                                                                                                                                                                                                                                                                                                                  1062
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1482
                                         642
                                                                                                                          762
                                                                                                                                                                                                          882
 582
                                                                                                                                                                  822
                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                 g
                                                                                ò
                                                                                                      g
                                                                                                                          ŏ
                                                                                                                                            g
                                                                                                                                                                οy
                                                                                                                                                                                   q
                                                                                                                                                                                                       δy
                                                                                                                                                                                                                                                ογ
                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                     Qγ
                                                                                                                                                                                                                                                                                                         Ω
                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                     οy
                                                                                                                                                                                                                                                                                                                                                                                                               δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4,
                                                                                                                                                                                                                                                                                                                                                                PAT 30-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                     Vertebrata; Euteleostomi;
                                                                                                                         actgggtcccccctcagtctttcagcaacctgggaccgaaccggtggcacaatgaacgga 1776
                                                                                                                                                                                                                                                ATTGGATCCCACCGCAGACACTGGCCACCTTCCCGGGTAGCTGGTGGAACCTGGAACGAA 1696
                                                                                                                                                                                    1754 GGGAGGACCTGGCTGGCCAGCGGCCTGATGTGGTCCGCACCCTGCTGGCTCGCCTGGCCG 1813
                                                                                                                                                                                                          agttcaacaaaactgcagtgccggtcaggtatccccccaaagaccccagaagtaacccta 1956
                                                                                                                                                                                                                                                                                                                                                                                                                         Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostom
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Dases 1 to 1710)
Glucksmann,M.A., Williamson,M., Rudolph-Owen,L.A. and Tsai,F.Y.
  tatacaccaaggcaaaaaatggctcctgggcagcaggctatgggatctggaacactgcaa 1656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gagoccagoacaacttocacotoccagocccatotoatttoatcotagoggatgatcag 581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             112 CAGCCCTCGGCCGCTCCGCCCCAGCCTCCATCTTCATCCTCCACGACGACGACAA 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  402 atgggggggcgctggcaggattctggatcctctgcctcctcacttatggttacctgtcctgg 461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                     gggtggacctatctaacaggtatccaggaatcgtgaagaagctcctacggaggctctcac
                                                                                                                                                                                                                             AATATAACCGCACAGCCATCCCGGTACGCTACCCAGCTGAGAACCCCCGGGCTCATCCTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ggccaggccttagaagaaggaagaagagggccttactagctcaagctggagagaaacta
                                         tccagtcagccatcagagtgcagcactggaaattgcttacaggaaatcctggctacagcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 1710;
                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 594.4; DB 6;
Pred, No. 1.7e-167;
); Mismatches 596;
                                                                                                                                                                                                                                                                                                                                                                DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Human sulfatases
Patent: WO 0155411-A 13 02-AUG-2001;
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                1710 bp
WO0155411.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1. 1710
/organism="Homo sapiens"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /db_xref="taxon:9606"
554 c 500 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 500 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                     Sequence 13 from Patent
AX206972
                                                                                                                                                                                                                                                                                                                                                                                            AX206972.1 GI:15394728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        / Match ' 26.2%;
Local Similarity 62.2%;
hes 1007; Conservative C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ø
                                                                                                                                                                                                                                                                                          gca 2019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Sim
Matches 1007;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                        DEFINITION
ACCESSION
                                                                                                                                                                                                                                                                                                                                            11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BASE COUNT
ORIGIN
                    1517
                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM
                                                                                                                                            1697
1597
                                        1657
                                                            1577
                                                                                 1717
                                                                                                      1637
                                                                                                                                                                                                          1897
                                                                                                                                                                                                                                                                                         2017
                                                                                                                                                                                                                                                                                                              1934
                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   522
                                                                                                                                                                 1837
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                      KEYWORDS
                                                                                                                                                                                                                                                                                                                                                    AX206972
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
                                                                                                                                                                                                                                                                                                                                                                                              VERSION
                                                                                                                                                                                   qq
                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                    유
                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                            엽
 à
                   g
                                       ô
                                                           g
                                                                               ò
                                                                                                   g
                                                                                                                         à
                                                                                                                                          a
                                                                                                                                                                 õ
                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

```
1301
                                                                                                                                                                                                                                                                                                                                                                                1121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1302 gataatggtggccagcctacggcaggagggagtaactggcctctcagaggtagcaaagga 1361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1069 crescassrestraccaccrescassesarisserrasarisserrasarisserrasarissers
                                                                                                                                                                                                                                                                                    tatacacactacaaatgtgacagtcctgggatgtgtggctatgacttgtatgaaaacgac 1001
                                                                                                                                                                                                                                      941
                                           701
                                                                   291
                                                                                          761
                                                                                                                  351
                                                                                                                                          821
                                                                                                                                                               411
                                                                                                                                                                                        881
                                                                                                                                                                                                              471
                                                                                                                                                                                                                                                            531
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      atgcccaccagaagaggatttgataccttttttggttcccttttgggaagtggggattac
                                                                                                                                                                                                                                                 aatgotgootgggaotatgacaatggcatataotocacacagatgtacactoagaggta
                                                                                                                                                                                                                                                                                                                                                                                               649 AGCCATATCCTGGCCACACAGCCTCAGCGTCCCCTCTTCCTCTATGTGGCCTTCCAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ctggctgaaggacagattgatgaggacattcaactagatggctatgatatctgggagacc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ataagtgagggtcttcgctcaccccgagtagatattttgcataacattgaccccatatac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATCAGCGAGGGCCGGGCCTCACCACGCACGGAGATCCTGCACAACATTGACCCACTCTAC
ggatttagagatgtgggttaccacggatctgagattaaaaacacctactcttgacaagctc
                                              gotgocogaaggagttaaactggagaactactatgtocagcotatttgcacaccatccagg
                                                                                                                                                                                                                                                                                                                                                 cagcaaatcttagcttcccataaccccacaaagcctatatttttatatattgcctatcaa
```

Wed Jul

g ô අ ò g g

à

g

ö g

ò

```
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local
                                                                                                                                                                               source
                                                                                                                                                                                                                                                                                                                              BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        960
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1080
                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Consensus quality: 180956 bases at least 040
Consensus quality: 180954 bases at least 030
Consensus quality: 181383 bases at least 030
Consensus quality: 181383 bases at least 030
Estimated insert size: 181600; pulse field gel estimation
Estimated insert size: 181623; sum-of-contigs estimation
Quality coverage: 7.04 in 020 bases; pulse field gel estimation
Quality coverage: 7.13 in 020 bases; pulse field gel estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AC011372 181923 bp DNA linear HTG 31-AUG-2001
Homo sapiens chromosome 5 clone CTB-108B20, WORKING DRAFT SEQUENCE,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Unpublished
2 (bases 1 to 181923)
Direct Submission
Submitted (06-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Malnut Creek, CA 94598, USA On Aug 25, 2001 this sequence version replaced gi:13699557.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Craniata, Vertebrata, Euteleostomi;
Catarrhini; Hominidae; Homo.
                                                                                                                                                                                              aacaaaactgcagtgccggtcaggtatcccccaaagaccccagaagtaaccctaggctc 1961
1662 tcagccatcagagtgcagcactggaaattgcttacaggaaatcctggctacagcgactgg 1721
                                                 1249 GCTGCCATCCGCGTGGGTGAGTGGAAGCTGCTGACAGGAGACCCCGGCTATGGCGATTGG 1308
                                                                                                   1722 gtccccctcagtctttcagcaacctgggaccgaaccggtggcacaatgaacggatcacc 1781
                                                                                                                                                                                                                                                                                                    gacctatctaacaggtatccaggaatcgtgaagaagctcctacggaggctctcacagttc 1901
                                                                                                                                                                                                                                                                                                                                                   GACCTGGCTGGCCAGCGGCCTGATGTGGTCCGCACCCTGCTGGCTCGCCTGGCCGAATAT 1485
                                                                                                                                                                                                                                                                                                                                                                                                                                                    1486 AACCGCACACCCATCCCGGTACGCTACCCAGCTGAGAACCCCCGGGCTCATCCTGACTTT 1545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             aatggaggggtctagggaccatggtatagagagaaaccaagaaaaagaagccaagca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11612: contig of 11612 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Project Information
Center Project Name: 105720, H226
Center clone name: CIT978SKB_108B20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Web site: http://www.jgi.doe.gov
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequencing of Human Chromosome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Center: Joint Genome Institute
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Chordata;
Mammalla; Eutheria; Primates;
1 (bases 1 to 181923)
DOE Joint Genome Institute.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AC011372.6 GI:15290338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----Genome Center
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    unordered pieces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Summary Statistics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Center Code: JGI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
                                                                                                                                                                                                 1782
                                                                                                                                                                                                                                                                                                                                                 1426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
REFERENCE
                                                                                                                                                                                                                                                   1369
                                                                                                                                                                                                                                                                                                                                                                                                  1902
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AC011372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
```

```
Db 101678 AGGTACCAGATCCACACAGGACTCCAGCATTCCATCATCGGCCCACAGCAGCCGAACTGC 101737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Db 102095 CAGTCCCTCGTGAGTACCTGTACCGCTACCGCACCATGGGCAATGTGGCCCGGCGGAAG 102154
                                                                                                                                                                                                                                                                                                                                                                                                                                                720 aagtatcagatacaccccggacttcaacattctatcataagacctacccaacccaactgt 779
                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 780 ttacctctggacaatgccaccctacctcagaaactgaaggaggttggatattcaacgcat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Db 101738 CTGCCCCTGGACCAGGAGCACACAGAAGCTGCAGGAGGCAGGTATATTCCACCAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Db 101798 AFGTGGGCAAGTGGCACCTGGGCTTCTACCGGAAGGAGTGTCTGCCCACCCGTCGGGGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             gacagtcctggggatgtgtgggctatgacttgtatgaaaacgacaatgctgcctgggactat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Db 101918 GATGGCCCAGGCGTGCGGCTTCGACCTGCAGGAGGGGGGAGAATGTGGCCTGGG---GG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1140 caageteetggeaggtatttegaacaetaecegateeattateaacataaacaggaggaga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Db 102215 CGCTACGGTTTCTACAACAACAGTGTCATCATCTTCTCCAGTGACAATGGTGGCCAGACT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1320 acggcaggagggagtaactggcctctcagaggtagcaaaggaacatattgggaaggaggg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      102275 TICTCGGGGGGGAGCAACTGGCCGCTCCGAGGACGCAAGGGCACTTATTGGGAAGGTGGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     900 tttgataccttttttggttcccttttgggaagtggggattactatacacactacaaatgt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1260 acttatggtttctataacaacagcattatcatttactcttcagataatggtggccagcct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1380 atccgggctgtaggctttgtgcatagcccacttctgaaaaacaagggaacagtgtgtaag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 181923:
                                                                                                                                                                                                                                                                                                                                                                                                       9
11712: gap of unknown length
41121: contig of 29409 bp in length
41221: gap of unknown length
75939: contig of 34718 bp in length
76039: gap of unknown length
181923: contig of 105884 bp in length.
                                                                                                                                                                                                                                                         human BAC library B"
44827 t 300 others
                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                           Score 496.8; DB 2;
Pred. No. 1.7e-137;
                                                                                                                                                                                                                                                                                                                                                                                                       482;
                                                                                                                                                                                                                                                                                                                                                                                                     0; Mismatches
                                                                                                                                                                      /organism="Homo sapiens"
/db_xref="taxon:9606"
                                                                                                                                                                                                                                 /clone="CTB-108B20"
/clone_lib="CalTech
44776 c 47833 g
                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                 /chromosome="5
                                                                                                                                                                                                                                                                                                                                                             21.9%;
62.5%;
                                                                                                                                                .181923
                                                                                                                                                                                                                                                                                                                                                                                                       Matches 812; Conservative
                                                                                                                                                                                                                                                                                                                                                                                   Similarity
  11613
11713
41122
41222
75940
                                                                                                                                                                                                                                                                               α
                                                                                                                                                                                                                                                                               44187
```

4

720

780

479

840

900

g

ò g õ

à

g

qq

ò g à

ð

```
GFHDVGYHGSEIHTPTLDALSASGVRLENYYVQPICTPTRSQLMSGRYQIHTGLQHGI
INSCQPNALPRUSPTLADKLKESGYATHMVGKWHLGFYKQEYLPWNRGFDTYFGYLNA
ABDYFHNHVWRQVRTLDLENDNGFWNRTGTGYSAHLFTGKATDVVQSHNTSKPLFLY
LAYOSWHAPLEVPERK EHKYRNITDKNRRFTRGMYSALLDEGYANLTQALKDKGLWNNY
VLIFSTDNGGQIHAGGNNYPLRGWKASLWEGGFHGVGFVSGGALKRSGAVSKGLIHVS
DWFPTLYTLAGGNLNGTKPLDGFNQMDTISNETPSPREILLHNIDILYPQKGYPLYSN
NFWDTRVRAAIRVGDYKLLITGDPGNSWYPPDGHLYFVPEIQESAAKNVWLFNITADP
NFWDTSSEKPLEVLELGLOVGFNYTAVPPRYPAPDPRCDFALHGDVWGPWE"
611 C 656 9 667 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               tccgggctgtaggctttgtgcatagcccacttctgaaaaaacaagggaacagtgtgtaagg 1440
                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                  cctcccagccccatctcattttcatcctagcggatgatcagggatttagagatgtgggtt
                                                                                                                                                                                                                                                                                                                                                     accacggatctgagattaaaacacctactcttgacaagctcgctgccgaaggagttaaaac
                                                                                                                                                                                                                                                                                                                                                                     tggagaactactatgtccagcctatttgcacaccatccaggagtcagtttattactggaa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           agtatcagatacacaccggacttcaacattctatcataagacctacccaacccaactgtt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           420 GATATCAGATACATACTGGGCTTCAACATGGTATTAATAATAATAGTTGTCAGCCAAACGCTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1141 aagctcctggcaggtatttcgaacactaccgatccattatcaacataaacaggaggagat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1201 atgctgccatgctttcctgcttagatgaagcaatcaacaacgtgacattggctctaaaga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TCGCGGGTATGGTCTCAGCCCTGGATGAGGGGGTAGCCAACTTAACCCAGGCACTCAAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          cttatggtttctataacaacagcattatcatttactcttcagataatggtggccagccta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACAAGGGACTCTGGAACAATACTGTCTTGATATTTTCTACAGAGAATGGCGGGGGGGAGATAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              cggcaggaggagtaactggcctctcagaggtagcaaaggaacatattgggaaggaggga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              360 IGGAGAACTACTACGIGCAGCCGATTIGCACCCCGACCAGGAGTCAGCTCATGTCCGGGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               tacctctggacaatgccacctacctcagaaactgaaggaggttggatattcaacgcata
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ttgataccttttttggttcccttttgggaagtggggattactatacacactacaaatgtg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      961 acagtcctgggatgtgtggctatgacttgtatgaaaacgacaatgctgcctgggactatg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      660 GGAGACAAGTG-----CGCTACCTTGACCTCAGAGATAACAACGGACCCGTGAGAAACG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30;
                                                                                                                                                                                                               Length
                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                            Score 398.2; DB 3;
Pred. No. 3.1e-108;
0; Mismatches 593;
                                                                                                                                                                                                            Query Match
Best Local Similarity 56.9%;
Matches 824; Conservative
                                                                                                                                           871
                                                                                                                                         BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1021
                                                                                                                                                                                                                                                                              541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    009
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1261
                                                                                                                                                                                                                                                                                                                                                   601
                                                                                                                                                                                                                                                                                                                                                                                  300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   901
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1081
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1321
                                                                                                                                                                                                                                                                                                                                                                                                                      661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             781
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INV 05-MAY-2000 complete cds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /translation-"MCKCLLVLIAIITACAVADQSSASAGTRQDAGQPNIVFVLADDF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GTACGCTACCCAGCTGAGACCCCCGGGCTCATCCTGACTTTAATGGGGGTGCTTGGGGG 102931
                                                                                                      102574
                                                                                                                                                                                                                                            102694
                                                                                                                                                                                                                                                                                                                                                                                    102811
                                                                                                                                                                                                                                                                                                                                                                                                                                                     CCTGATGTGGTCCGCACCCTGCTCGCTCGCCGGAATATAACCGCACAGCCATCCCG 102871
                                                                                                                                         1679
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1979
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Helix ponetta

Bukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata;

Stylomnatophora; Helicacea; Helicidae; Helix.

1 (bases 1 to 280);

Wittstock,U., Fischer,M., Svendsen,I. and Halkier,B.A.

Cloning and characterization of two cDNAs encoding sulfatases

the Roman snail, Helix pomatia

100331298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Submitted (27-NOV-1998) Plant Biology, The Royal Veterinary a Submitted University, Thorvaldsensvej 40, Frederiksberg C 1871, Denmark
                                                                                                                                                                                                                              102635 GAGTGGAAGCTGCTGACAGGACCCCGGCTATGGCGATTGGATCCCACCGCAGACTG
                                                                                                                                                                                                                                                                                                                                                 gatgaggacattcaactagatggctatgatatctgggagaccataagtgaggtcttcgc
                     tcctgggcagcaggctatgggatctggaacactgcaatccagtcagccatcagagtgcag
                                                                                                                                                             cactggaaattgcttacaggaaatcctggctacagcgactgggtcccccctcagtcttc
                                                                                                                                                                                                                                                                              agcaacctgggaccgatggcacaatgaacggatcacctcgtcaactggcaaagt
                                                                                                                                                                                                                                                                                                             102695 GCCACCTTCCCGGGTAGCTGGTGGAACCTGGAACGAATGGCCAGTG---TCCGCCAGGCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         gtcaggtatcccccaaagaccccagaagtaaccctaggctcaatggaggggtctaggga
                                                                      tcaccccgagtagatattttgcataacattgaccccatatacaccaaggcaaaaatggc
                                                                                                                                                                                                                                                                                                                                                                                                                      ccaggaatcgtgaagaagctcctacggaggctctcacagttcaacaaaactgcagtgccg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 (bases 1 to 2805)
Wittstock,U., Fischer,M., Svendsen,I. and Halkier,B.A.
Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Helix pomatia sulfatase 1 precursor (SULF1) mRNA, AF109924 1 GI:6863175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     102932 CCCTGGGCCAGTGAGGAAGAGGAGGAAGAGGAA 102971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ccatggtatagagagaaaccaagaaaaagaagccaagca 2019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /codon_start=1
/product="sulfatase 1 precursor"
/protein_id="AAF30402.1"
/db_xref="GI:6863176"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /db_xref="taxon:6536"
11. .2805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers
1. .2805
/organism="Helix pon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   149. .1660
/gene="SULF1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /gene="SULF1"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Roman snail
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13
                                                                                                                                       1620
                                                                                                                                                                                                                                                                             1740
 1500
                                                                    1560
                                                                                                                                                                        102575
                                                                                                                                                                                                          1680
                                                                                                                                                                                                                                                                                                                                                                                                                    1860
                                                                                                                                                                                                                                                                                                                                                                                                                                                     102812
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         102872
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 1
AF109924
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CDS
```

g

ò

g

ò g à 1020

629

1080

773

Wed

음 ò g ó g g

ò

셤

õ

g

ö

셤 ò g ò 셤 ò

ò

```
CEANGE AT A LICENSE AND SERROLD SPREADLE STRUCTURE SERVED 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         966 cctggga------tgtgtgggctatgacttgtatgaaaacgacaatgctgcc 1010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1011 tgggactatgacaatggcatatactccacacagatgtacactcagagagtacagcaaatc 1070
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1191 aggaggagatatgctgccatgcttcctgcttagatgaagcaatcaacaacgtgacattg 1250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             725
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    accttttttggttcccttttgggaagtggggattactatacacactacaaatgtgacagt 965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AGGCATTACTATGCAGGAATGGTGTCTCTTATGGATGAAGCAGTGGGAAATGTCACAGCA 855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                546 cagccccatctcattttcatcctagcggatgatcaggggatttagagatgtgggttaccac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ggatctgagattaaaacacctactcttgacaagctcgctgccgaaggagttaaactggag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           aactactatgtccagcctatttgcacaccatccaggagtcagtttattactggaaagtat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      736 GAACCCCTTCAGGTCCCTGAGGAGTACCTGAAACCCTACGACTTTATCCAAGATAAGAAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            cagatacacaccggacttcaacattctatcataagacctacccaaccgactgtttacct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          319 CAGATCCACACAGGTTTACAACACCAAATAATCTGGCCCTGTCAGCCCAGCTGCGTCCCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ctggacaatgccaccctacctcagaaactgaaggaggttggatattcaacgcatatggtc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        619 ACAGGATATAAAAAT---ATGTATTCGACAAACATATTACTGAAAGAGCTACAGCCCTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99
                                                  2; ARSB'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 1939;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                  Fig.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16.7%; Score 379; DB 4; I
56.4%; Pred. No. 1.7e-102;
ive 0; Mismatches 575;
                                                    from
                           B, ARSB"
                    /gene="arylsulfatase B, ARS
/note="This sequence comes
/codon_start=1
                                                                                             /product="arylsulfatase B"
/protein_id="AAB23941.1"
/db_xref="GI:258856"
                                                                                                                                                                                                                                                                                                                                                                                                                503 g
                                                                                                                                                                                                                                                                                                                                                                                                                     552 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                     450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local
                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    906
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       984
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       439
     CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MAM 08-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               feline
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Jackson, C.E., Yuhki, N., Desnick, R.J., Haskins, M.E., O'Brien, S.J. and Schuchman, E.H.
Felline arylsulfatase B (ARSB): isolation and expression of the cDNA, comparison with human ARSB, and gene localization to fellin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           cctgggcagcagctatgggatctggaacactgcaatccagtcagccatcagagtgcagc 1680
                                                                                                                                                                                                                                                                                                                                                                                                                                                                1590 CTAGATACCCCCCAGACCCCAGGTGCGACCCAGCTCTGCATGGTGATGTCTGGGGAC 1649
                                                                                                  atgaggacattcaactagatggctatgatatctgggagaccataagtgagggtcttcgct 1560
                                                                                                                                                                                                                                                                   caccccgagtagatattttgcataacattgaccccatatacaccaaggcaaaaatggct 1620
                                                                                                                                                                                                                                                                                          CCTGTACAGCA-----ACACCTGGGACACGAGGTCAGGGCAGCTATCAGAGTTGGGG 1364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gcaacctgggaccgaaccggtggcacaatgaacggatcacctcgtcaactggcaaagtg 1800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ------CACCTGTATTTTGTACTGAAATCCAAGAATCGGCTGCGAAAAACG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1921 tcaggtatcccccaaagaccccagaagtaaccctaggctcaatggagggtctagggac 1980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 117975] from the original journal article. This sequence comes from Fig. 2.
                                                                         aacctgtgcacatcactgactggtaccccactctcatttcactggctgaaggacagattg
                                                                                                                                                                                                    caggaatcgtgaagaagctcctacggaggctctcacagttcaacaaaactgcagtgccgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelec
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
1 (bases 1 to 1939)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         S48472 1939 bp mRNA lin
arylsulfatase B [cats, liver, mRNA, 1939 nt].
848472.1 GI:258855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /gene="arylsulfatase B, ARSB"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /db_xref="taxon:9685"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Genomics 14 (2), 403-411 (1992) 93052342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             .1939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .1939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  location: A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          chromosome Al
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Felis catus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1981 catggta 1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1650 CGTGGGA 1656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  cat liver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM
                                                                      1441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                1312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1530
                                                                                                                   1134
                                                                                                                                                                   1501
                                                                                                                                                                                                                   1194
                                                                                                                                                                                                                                                                                                                1254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1741
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1861
                                                                                                                                                                                                                                                                      1561
                                                                                                                                                                                                                                                                                                                                                                   1621
                                                                                                                                                                                                                                                                                                                                                                                                                                                              1681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

셤

셤

٠.

```
/pocesi__de__aaa51779.1"
/db_xref_=GI:179030"
/db_xref_=GB:G00-119-008"
/translation="worrdaascpreedprrilepvvlplllilllappgsgaggrp
PHLVFLLADDLGWNDVGFHGSRIRTPHLDALAAGGVLLDNYYTOPLCTPSRSQLLTGR
YORRGLQMQTIWPCOPSCYPLDEKLLPQLKEAGYTTHWOKWHLGWYRKECLPTRR
YORRGLQMQTIWPCOPSCYSTBECTLIDALWYRKAGLPFRR
AIALITNHPPERPLFLYSHEQTLIDALWYRKAGLPFRRAGAGNTAGSBOYTSHWAGWVSTNIFTKR
AIALITNHPPERPLFLYLALQSVHEPLQVPEEYLKPYDFIODKNRHHYAGWVSLMDER
                                                                                                                                                                                                                                                                                                                                VGNVTAALKSSGLWNNTVFIFSTDNGGQTLAGGNNWPLRGRKWSLWEGGVRGVGFVAS
PLEQNGGVRNRELIH 15DWLPFLVKLRGHTWGYRFLDGFDWKTISEGSPSPRIELL
HNIDPNFVDSSPCPHYMAPARAPEDSELPEY SAFNTSVHAAIRHGNWKLLTGYPGCGYW
FPPPSQYNVSEIPSSDPPTKTLWLFDIDRDPERHDLSREYPHIVTKLLSRLQFYKH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  981 ATATGGTCGGAAAATGGCACCTGGGAATGTACCGGAAAGAATGCCTTCCAACCCGCCGAG 1040
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GTACATTAATTGACGCTCTGAATGTCACACGATGTGCTCTTGATTTTCGAGATGGCGAAG 1160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1003 atgotgoctgggactatgacaatggcatatactccacacagatgtacactcagagagtac 1062
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ccacctcccagccccatctcattctagcggatgatcagggatttagagatgtgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gttaccacggatctgagattaaaacacctactcttgacaagctcgctgccgaaggagtta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 921 GIGITCCTCTGGATGAAAACTCCTGCCCCAGCTCCTAAAAGAAGCAGGTTATACTACCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    aactggagaactactatgtccagcctatttgcacaccatccaggagtcagtttattactg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              801 recregaciaciaciacacecaececierecaeecereeeegaeecaecreere
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gaaagtatcagatacacccggacttcaacattctatcataagacctacccaacc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gtttacctctggacaatgccaccctacctcagaaactgaaggaggttggatattcaacgc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g-----tgacagtcctgggatgtgtggctatgacttgtatgaaaacgaca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GCCGCTACCAGATCCGTACAGGTTTACAGCACCAAATAATCTGGCCCTGTCAGCCCAGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 2802;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                 /note≕"arylsulfatase B signal peptide'
880. .2158
                                                                                                                                                                   precursor"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 16.2%; Score 368.2; DB 9; Best Local Similarity 55.8%; Pred. No. 3.5e-99; Matches 827; Conservative 0; Mismatches 588;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 е в<sup>в</sup>
                                                                                                                                                                                                                                                                                                                                                                                                   SVPVYFPAQDPRCDPKATGVWGPWM"
/db_xref="taxon:9606"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /product="arylsulfatase
767 c 726 g 60
                                                                                                                                                                 /note="arylsulfatase B
                                                                               /product="ASB mRNA"
                              /map="5p11-q13"
                                                                                                                                                                                  /codon_start=1
                                                                 /gene="ARSB"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /gene="ARSB"
                                                                                                                   /qene="ARSB"
                                                                                                                                 560. .2161
/gene="ARSB"
                                                                                                                                                                                                                                                                                                                                                                                                                560. .679
/gene="ARSB"
                                                                                                  .2802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  646
                                                                                                                                                                                                                                                                                                                                                                                                                  sig_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                538 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  178
                                                                                                   gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 658
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 861
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    838
                                                 mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1101
                                                                                                                                 CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRI 31-0CT-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Euteleostomi;
                                                                                                                               gtgtgtaaggaacctgtgcacatcactgactggtaccccactctcatttcactggctgaa 1490
                                                                                                                                                                                                                                                                TAGCCTTCAACACATCTCTGCATGCTGCAATTAGACACGGGAACTGGAAACTCCTCACC 1335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1336 GGCTACCCAGGCTGTGGCTGTTGCTTCCGCCGTCTCCATACAACGATTCTGCGATA 1395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1441 GATCAGGACCCAGAAGAAAGACATGACCTGTCAAGAGCTATCCCCCATATTGTCGAGCAG 1500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gctctaaagacttatggtttctataacaacagcattatcatttactcttcagataatggt 1310
                                                                                                                                                                                                                                                                                             1096 GGAAGCACAAAAGGCACCAAGCCTCTGGATGGCTTTGATGTGGGAAAACCATCAGTGAA 1155
                                                                                                                                                                                                                                                                                                                                                                                                                              1216 TCACCGTGTCCTGGGAAAAGCCTGGCACCAGCAAGGATGATTCTTCTTCATCCAGCGTAT 1275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gggatctggaacactgcaatccagtcagccatcagagtgcagcactggaaattgcttaca 1697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ggaaatcctggctacagcgactgggtccccctcagtcttcagcaacctgggaccgaac 1757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                cggtggcacaatgaacggatcacctcgtcaactggcaaaagtgtatggcttttcaacatc 1817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              acagecgaeccatatgagagggtggaectatetaacaggtatecaggaategtgaagaag 1877
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     oţ
                  ggccagcctacggcaggaggagtaactggcctctcagaggtagcaaaggaacatattgg
                                                                                -----acaccaaggcaaaaatggctcctgggcagcagctat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       scnuchman, E.H., Jackson, C.E. and Desnick, R.J.
Human arylsulfatase B: MOPAC cloning, nucleotide sequence full-length cDNA, and regions of amino acid identity with arylsulfatases A and C
Genomics 6 (1), 149-158 (1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; E
Mammalla; Eutheria; Primates; Catarrhini; Hominidae;
11 (bases 1 to 2802)
Schuchman, E.H., Jackson, C.E. and Desnick, R.J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      томакъых
Human arylsulfatase B (ASB) mRNA, complete cds.
M32373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             arylsulfatase; lysosomal hydrolase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Location/Qualifiers
1. .2802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Human liver, cDNA to mRNA.
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           M32373.1 GI:179029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 15
HUMARSBX
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    source
                                                                                                                                                                                                                            1036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1818
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM
                                                                                                                             1371
                                                                                                                                                                                                                                                                                                                                                                                               1600
                                                                                                                                                                                                                                                                                                                                                                                                                                                               1638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1878
                                                                                                                                                                                              1431
                                                                                                                                                                                                                                                              1491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE
FEATURES
                                                               1311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                JOURNAL
                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
ò
                              쉽
                                                               õ
                                                                                            g
                                                                                                                             ò
                                                                                                                                                            q
                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
```

٠,

Gaps

99

657

717 960 777 837 980 897

$\vdash$	1122 1277	1182 1337	242 397	1302 1457	362 517	422 577	482 637	542 697	1601 1757	629 817	1689 1877	1749 1937	1809 1982	1869 . 2042	1929 2102	
61 AAGTTGCAACAGGATATAAAAATATGTATTCAACAAAACATATTCACCAAAAGGGCTA	1063 agcamatcttagcttcccataaccccacaaagcctatattttatatattgcctatcaag 11	1123 ctgttcattcaccactgcaagctcctggcaggtatttcgaacactaccgatccattatca 11	1183 acataaacaggaggagatatgctgccatgcttcctgcttagatgaagcaatcaacaacg 12   1   1   1   1   1   1   1   1   1	1243 tgacattggctctaaagacttatggtttctataacaacagcattatcatttactcttcag 13   1   1   1   1   1   1   1   1   1	1303 ataatggtggccagcctacggcaggaggagtaactggcctctcagaggtagcaaaggaa 1 	1363 catattgggaagggagtccgggctgtaggctttgtgcatagcccacttctgaaaaca 14   1   1   1   1   1   1   1   1   1	1423 agggaacagtgtgtaaggaacctgtgcacatcactgactg	1483 tggctgaaggacagattgatgaggacattcaactagatggctatgatatctgggagacca 1   1   1   1   1   1   1   1   1   1	1543 taagtgagggtcttcgctcaccccgagtagatattttgcataacattgaccccatatac-1	1602	1630 caggctatgggatctggaacactgcaatccagtcagccatcagagtgcagcactggaaat 1   1   1   1   1   1   1   1   1   1	1690 tgcttacaggaaatcctggctacagcgactgggtccccctcagtctttcagcaacctgg 1	1750 gaccgaaccggtggcacaatgaacggatcacctcgtcaactggcaaaagtgtatggcttt 1 	1810 tcaacatcacagccgacccatatgagaggtggacctatctaacaggtatccaggaatcg 1	1870 tgaagaagctcctacggaggctctcacagttcaacaaactgcagtgccggtcaggtatc 1	1930 ccccaaagaccccagaagtaaccctaggctcaatggaggg 1970 
qq	oy Ob	Oy Dp	Qy Db	oy Dp	Oy Dp	Qy Dp	Qy	QY Db	Qy	Qy	Qy Db	Oy Dp	Oy Db	Oy Dp	Oy Dp	Qy Dp

Search completed: July 23, 2002, 17:14:59 Job time: 7805 sec

	<i>A</i>		
**			
÷.			

-

OM nucleic

Run on:

```
AA123795 mp93a07.r
A1186106 mp93c07.y
AA124434 mp93c07.r
BC346926 dad19f07.
BC194726 RST13892
AA159952 zo75a09.r
AA159952 zo75a09.r
AA137376 mp93c07.x
B1441091 ic80903.y
A738292 AV38292.
BF160020 601767706
AA148125 mo62a02.x
BB654214 BB654214
AA117137 mo52a02.x
BB654214 BB654214
AA117137 mo62a02.x
BB654214 BB654214
AA117137 mo62a02.x
BB633888 BB633888
BB014208 125817 MAR
BB633888 BB633888
BB033888 BB633888
BB033888 BB633888
                                                                                                                                                                                                                                                                                        BE163345 601771896
AL666493 AL666493
AI815689 au49f04.y
BE305390 601099974
BF218442 601881811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BG680752 817 bp mRNA linear EST 01-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                 AI791623 nj45c11.y
BE786803 601477606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /tissue_type="squamous cell carcinoma"
/lab_host="DH108 (TI phage-resistant)."
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: Sall; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4753259"
/clone_lib="NCI_CGAP_Skn4"
                                                                                                                                                                                                                                                                                                                                                                                                                        ALIGNMENTS
                                                                                                                                                    AV382922
BF160020
AT448125
BB654214
AA117137
AW358659
BB614288
BB614208
BC14208
BF69654
BB6087596
                                                                                                                                                                                                                                                                                           BF163345
AL666493
AI815689
BE305390
                                                                                            AU055665
AU055666
AI327376
BI441091
                                                                                                                                                                                                                                                                                                                                                   BF218442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BG680752
BG680752.1 GI:13912149
                          0000000
                                                                                                                                                                                                                                                                                                                      100
1. .817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ø
 000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             human.
241.4
226.6
226.6
209.1
199.4
198.6
198.6
198.6
198.6
196.2
196.2
196.2
196.2
196.2
197.8
198.8
116.9
116.9
116.9
116.1
116.9
116.1
116.1
116.1
116.2
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
116.3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
BG680752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURES
 O
                                                                                                 000
                                                                                                                                                                                                              υ
                                                                                                            (without alignments)
12773.446 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AW474222 xs20h05.x
AZ815176 ZM0083G13
BB520951 BB520951
BB319509 BB319509
AA159953 ZG72009.s
W01455 Zd87204.s1
BF653682 Z77595 MA
BG661302 ia59e07.y
W01485 Zd87204.r1
R43179 y918003.s1
U46280 HSU46280 Hu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BM461300 AGENCOURT
BE553565 ur47a08.y
BI104707 602891618
BF979627 602287919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BG680752 602628445
BF307503 601893984
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                               2266
1 cacgcgtccgcccacgcgtc.....tgccacctggtgccgaattc 2266
                                                                                          July 23, 2002, 14:59:59; Search time 2394.35 Seconds
                                                                                                                                                                                                                                                                               27472414
                           Compugen Ltd
                                                                                                                                                                                                                                                 13736207 seqs, 6748477542 residues
                                                                                                                                                                                                                                                                         Potal number of hits satisfying chosen parameters:
           version 4.5
- 2000 Comp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUMMARIES
                                                                                                                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BGG80752
AW74203
AZ815175
BB520951
BB520951
BB119509
AA159953
W81455
BGG61302
W81485
BGG1302
W81485
BGG1302
W81485
BGG1302
W81485
BGG1302
W81485
BGG1302
W81485
BGG1302
W81485
                                                                  nucleic search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BI104707
BF979627
                                                                                                                                                                                                        IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
             GenCore
Copyright (c) 1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      em_gss_inv:*
em_gss_pln:*
em_gss_vrt:*
                                                                                                                                                                                                                                                                                                        DB seq length: 0
DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gb_gss:*
em_gss_hum:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0000000000
                                                                                                                                                    US-09-495-823-8
                                                                                                                                                                                                                                                                                                                                                                                                                                   em_esthum:*
em_estin:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                em_estmu:*
em_estov:*
em_estpl:*
em_estro:*
em_htc:*
gb_estl:*
gb_est2:*
gb_est2:*
                                                                                                                                                                                                                                                                                                                                                                                                                     em_estba:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                615
625
648
700
463
414
547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           458
325
1157
628
628
1032
                                                                                                                                                                                                                                                                                                                                                                                                           EST: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10:
                                                                                                                                                    Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   646.2
603.8
512.4
511.2
5511.2
5506.6
368.8
361.4
341.8
333.2
384.6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 261.4
245.8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score
                                                                                                                                                                                                           Scoring table:
```

••

Database

O

O

Result õ

Searched:

Minimum Maximum

```
203
                                                                                                                                               source
                                                                                                                                                                                                                                                                                                   COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              302 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1012
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                          592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         952
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          832
                                                                                                                                                                                                                                                                                                   BASE CO
                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 òχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BF307503 683 bp mRNA linear EST 21-NOV-2000 601893984F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4139565 5',
                                                   ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens
Bukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
1 (bases 1 to 683)
                                                                                                                                                                                                                                                                              1727
                                                                                                                                                                                                                                                                                                                      1787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ggggtctagggaccatggtatagaggaaaccaagaaaaaaagaagcaaagcaaaatcag 2027
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gotgagaaaaagcaaaagaaaago--aaaaaaagaagaagaaacagcagaaagcagtoto 2085
                                                                                                               1428 acagtgtgtaaggaacctgtgcacatcactgactggtaccccactctcatttcactggct 1487
                                                                                                                                                      gaaggacagattgatgaggacattcaactagatggctatgatatctgggagaccataagt 1547
                                                                                                                                                                                               gagggtcttcgctcaccccgagtagatattttgcataacattgaccccatatacaccaag 1607
                                                                                                                                                                                                                                      gcaaaaaatggctcctgggcagcaggctatgggatctggaacactgcaatccagtcagcc 1667
                                                                                                                                                                                                                                                                                                                                                             actggcaaaagtgtatggcttttcaacatcacagccgacccatatgagagggtggaccta 1847
                                                                                                                                                                                                                                                                                                                                                                                                      tctaacaggtatccaggaatcgtgaagaagctcctacggaggctctcacagttcaacaaa 1907
                                                                                                                                                                                                                                                                                                                                                                                                                                              actgcagtgccggtcaggtatcccccaaagaccccagaagtaaccctaggctcaatgga 1967
                                                                      1368 tgggaaggaggatccgggctgtaggctttgtgcatagcccacttctgaaaaacaaggga 1427
                                                                                                                                                                                                                                                                             608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          899
                                                                                                                         189 GAGGGTCTTCGCTCACCCCGAGTAGATATTTGCATAACATTGACCCCATATACACCAAG
                                                                                                                                                                                                                                                                                                                                                                                                                 GCAAAAAATGGCTCCTGGGCAGCAGCTATGGGATCTGGAACACTGCAATCCAGTCAGCC
                                                  ;
7
                               Length
                                                   Indels
                               DB 10;
                             Score 703.8; DB 10.
Pred. No. 3.1e-125;
); Mismatches 12;
                                                  ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BF307503.1 GI:11254633
                            Query Match 31.1%;
Best Local Similarity 98.1%;
Matches 723; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     aggiticaacitgccatt 2102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  mRNA sequence.
BF307503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         human.
                                                                                                                                                                          129 (
                                                                                                                                                                                                                                                          249
                                                                                                                                                                                                                                                                                                 309
                                                                                                                                                                                                                                                                                                                                                                                                                                           1908
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                                                                          6
                                                                                                                                                     1488
                                                                                                                                                                                             1548
                                                                                                                                                                                                                                      1608
                                                                                                                                                                                                                                                                                                                                                                                  429
                                                                                                                                                                                                                                                                                                                                                                                                      1848
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM
                                                                                                                                  69
                                                                                                                                                                                                                                                                             1668
                                                                                                                                                                                                                                                                                                                                         369
                                                                                                                                                                                                                                                                                                                                                             1788
                                                                                                                                                                                                                                                                                                                                                                                                                         489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         729
                                                                                                                                                                                                                                                                                                                     1728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
BF307503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
 ORIGIN
                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dp
                                                                                          q
                                                                                                               ò
                                                                                                                                  a
                                                                                                                                                     ö
                                                                                                                                                                         g
                                                                                                                                                                                             ò
                                                                                                                                                                                                               g
                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
```

```
/tissuc_type="Inbodomyosarcoma" / Tissuc_type="Inbodomyosarcoma" / Tissuc_type="Inbodomyosarcoma" / Tissuc_type="Inbodomyosarcoma" / Tab_host="DH108 (phage-resistant)" / Anote="Organ: muscle; Vector: pOTB7; Site_1: EcoRI; Site_2: Xho1; CDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size=selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ä
                                                                                                                                    CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1050 row: g column: 22
High quality sequence stop: 679.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1011
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        891
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               951
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ccaactgtttacctctggacaatgccaccttacctcagaaactgaaggaggttggatatt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    acaaatgtgacagtcctgggatgtgtggctatgacttgtatgaaaaacgacaatgctgcct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ttactggaaagtatcagatacacaccggacttcaacattctatcataagacctacccaac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gaagaggatttgatacctttttggttcccttttgggaagtggggattactatacacact
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gagttaaaactggagaactactatgtccagcctatttgcacaccatccaggagtcagttta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACAAATGTGACAGTCCTGGGATGTGTGGCTTTGTATGAAAACGACAATGCTGCCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 683;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 646.2; DB 10.
Pred. No. 3.7e-114;
0; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                  /organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4139565"
/clone_lib="NIH_MGC_17"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 28.5%;
Best Local Similarity 99.1%;
Matches 671; Conservative (
```

Wed Jul

```
AZ815176
AZ815176.1 GI:12985084
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      aagcaaaagaaaagc 2051
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15 AAGCAAAAGAAAGC 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               house mouse.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
VERSION
                                                                                                                     1677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2037
                                                   1617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 4
AZ815176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMMENT
                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                     Óγ
                                                                                                                                                  qq
                                                                                                                                                                                         ò
                                                                                                                                                                                                                           염
                                                                                                                                                                                                                                                         δy
                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                            Qγ
                                                                                                                                                                                                                                                                                                                                                            PP
                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1. .615
//organism="Homo sapiens"
//db_xref="taxon:9606"
//db_xref="taxon:9606"
//dlone=lib="NCI_CGAP_Ut2"
//clone=lib="NCI_CGAP_Ut2"
//tissue_type="moderately-differentiated endometrial
adenocarcinoma, 3 pooled tumors;
//lab_host="DH108"
//lab_host="DH108"
//lab_host="DH108"
//lab_host="Organ: uterus; Vector: pCMV-SPORT6; Site_1: Sall;
Site_2: Not1; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.85 kb. Life Technologies catalog #:
11539-012"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
WWW-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 435.
Location/Qualifiers
                                                                                                                                                                                                                                                     AW474222 615 bp mRNA linear EST 24-FEB-2000 xs20h05.xl NCI_CGAP_Ut2 Homo sapiens cDNA clone IMAGE:2770233 3' similar to SW:ARSB_HUMAN P15848 ARXLSULFATASE B PRECURSOR;, mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ő
                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 615)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1131 tcaccactgcaagctcctggcaggtatttcgaacactaccgatccattatcaacataaac 1190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1437 aaggaacctgtgcacatcactgactggtaccccactctcatttcactggctgaaggacag 1496
1071 ttagettcccataaccccacaaagcctatatttttatatattgcctatcaagctgttcat 1130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1557 cgctcaccccgagtagatattttgcataacattgaccccatatacaccaaggcaaaaat 1616
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Email: cgapbs_r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 603.8; DB 9; Length 615;
Pred. No. 5.3e-106;
0; Mismatches 7; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     б
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   153
                                                                                                                                                                                                                                                                                                         sequence.
AW474222
AW474222.1 GI:7044328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     26.6%;
                                                                                                                                   1191 aggaggagatatgctgc 1207
                                                                                                                                                     660 AGGAGGAGATATGCTGC 676
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 151 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
Matches 608; Conserv
                                                                                                                                                                                                                                                                                                                                                                              human.
                                                                                                                                                                                                                                                                                                                                                            EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                     RESULT 3
AW474222/c
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BASE COUNT
                                                                                                                                                                                                                                                                                                                         ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1497
                                                                                                                                                                                                                                                                                                                                                           KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORIGIN
                                g
                                                                 ŏ
                                                                                                 g
                                                                                                                                   ò
                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
```

```
E 1 (bases 1 to 625)
S Dun, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AZ815176 625 bp DNA linear GSS 20-FEB-2001
2M0083G13F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0083G13 F, DNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ΙŢ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                                         1736
                                                                                                                                                                                                                                                                                              1797 agtgtatggcttttcaacatcacagccgacccatatgagagggtggacctatctaacagg 1856
                                                                                                                                                                                                                                                                                                                                                                                              1857 tatccaggaatcgtgaagaagctcctacggaggctctcacagttcaacaaaactgcagtg 1916
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1917 ccggtcaggtatcccccaaagaccccagaagtaaccctaggctcaatggagggtctag 1976
                                                                                                                                                                                                                                                                                                                                             196
                                                                                                                                                                                                                                                                                                                                                                                                                        9/
435 GGCTCCTGGGCAGCAGCTATGGGATCTGGAACACTGCAATCCAGTCAGCCATCAGAGTG
                                                                                            135 CCGGTCAGGTATCCCCCCCAAAGACCCCAGAAGTAACCCTAGGCTCAATGGAGGGGGTCTGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Plate: 0083 row: G column: 13
Seq primer: CGTTGTAAAACGACGGCCAGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /organism="Mus musculus"
/strain="C57BL/6J"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Class: plasmid ends
High quality sequence stop: 625.
Location/Qualifiers
```

616 GGGTGGACCT 625

```
// And the control strain All Colu, in Testicate, F. // And Colure - Vector: PWA272V; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.bax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end repaired with T4 DNA polymerase and T4 polymucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pwD42 (qil4732114 qib) lat129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the lisaert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          actgggtccccctcagtcttcagcaacctgggaccgaaccggtggcacaatgaacgga 1776
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                aaggaacatattgggaaggagggatccggggctgtaggctttgtgcatagcccacttctga 1416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       tatacaccaaggcaaaaaatggctcctgggcagcagctatgggatctggaacactgcaa 1656
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             tccagtcagccatcagagtgcagcactggaaattgcttacaggaaatcctggctacagcg 1716
                                                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        cttcagataatggtggccagcctacggcaggaggagtaactggcctctcagaggtagca 1356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
/clone="UUGC2M0083G13"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22.6%; Score 512.4; DB 12; Length 625; 90.0%; Pred. No. 1.8e-88; 1.ve 0; Mismatches 61; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best_Local Similarity 90.0
Matches 549; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gggtggacct 1846
                                                                                                                                                                                                                                                                                                                                                                                                                         179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1657
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
```

```
Eukaryotta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostoml;

Bukaryotta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostoml;

Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 648)

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.

Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda

M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,

Okazaki, Y., Okido, T., Saito, R., Sakai, K., Sano, H., Sasaki

D., Shibata, K., Sainagawa, A., Shiraki, T., Sogabe, Y., Sano, H., Sasaki

Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,

Muramatsu, M. and Hayashizaki, T., et al. 2001)

On Jul 28, 2000 this sequence version replaced gi:9572409.

Contact: Yoshihde Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9216
                                                              BB520951 BIKEN full-length enriched, 16 days neonate heart Mus mussolus cDNA clone D830047F08 3', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. .10 (11), 1757-1771 (2000) Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawal,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAS Compared with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Email: genome_resegsc.riken.go.jp,
URL:http://genome_gsc.riken.go.jp,
Carninci.p., Shibata,Y., Hayasusu,N., Sugahara,Y., Shibata,K., Itoh
A., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
penes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yonedd,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
Hayashizaki,Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /clone_lib="RIKEN full-length enriched, 16 days neonate heart"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/)
further details.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /organism="Mus musculus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /tissue_type="heart"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Cocation/Qualifiers
                                                                                                                                                                                                                                   BB520951.2 GI:16443460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             prepare mouse tissues
                                                                                                                                                                                                                                                                                                                                                                Mus musculus
                                                                                                                                                                                                                                                                                                                      house mouse.
                                                                                                           DEFINITION
                                                                                                                                                                                                                                                                                                                                                                ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
AUTHORS
                                                                                                                                                                                                  ACCESSION
                                                                                                                                                                                                                               VERSION
KEYWORDS
SOURCE
                                     BB520951
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMMENT
RESULT
```

```
house mouse.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 source
DEFINITION
                                                  ORGANISM
                 ACCESSION
                        VERSION
KEYWORDS
                                                                          REFERENCE
                                                                                  AUTHORS
                                                                                                                                                   JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                           TITLE
                                                                                                                                                          COMMENT
                                         SOURCE
                    ;
                                                                                                                                                                                                           cagaggtagcaaaggaacatattgggaaggagggatccggggctgtaggctttgtgcatag 1405
                                                                                                                                                                                                                                          acccatatgagagggtggacctatctaacaggtatccaggaatcgtgaagaagtcctac 1884
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ggaggeteteacagtteaacaaaactgeagtgeeggteaggtateeeceeaaagaeeeca 1944
                                                                                                                                                                                                                                                                                                                                                                             ggaacactgcaatccagtcagccatcagagtgcagcactggaaattgcttacaggaaatc 1704
                                                                                                                                                                                                                                                                                                                                                                                                             ctggctacagcgactgggtccccctcagtctttcagcaacctgggaccgaaccggtggc 1764
                                                                                                                                                                                                                                                                                                                                           1585 acattgaccccatatacaccaaggcaaaaatggctcctgggcagcaggttatgggatct 1644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                540
                                                                                                                                                                                                                                                                                                                                                    541 GGAGACTGTCACAATTCAACAAGACTGCCGTGCCTGTCAGGTACCCCCGGAAGGATCCCA 600
                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                   /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                                                                                                                                                                                                                                                                                                                                                                                                                      1;
                                                                                                                                                                         22.6%; Score 511.2; DB 9; Length 648; llarity 87.8%; Pred. No. 3.1e-88; Conservative 0; Mismatches 78; Indels 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /dev_stage="16 days neonate"
/lab_host="DH10B"
                                                                                                                                                                                  Similarity
                                                                                                                                         179
                                                                                                                                                                                          Matches 569;
                                                                                                                                                                          Query Match
Best Local
                                                                                                                                        BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                            301
                                                                                                                                                                                                          1346
                                                                                                                                                                                                                                          1406
                                                                                                                                                                                                                                                                           1465
                                                                                                                                                                                                                                                                                                                                                                            1645
                                                                                                                                                                                                                                                                                                                                                                                                            1705
                                                                                                                                                                                                                                                                                                                                                                                                                                             1765
                                                                                                                                                                                                                                                                                                                                                                                                                                                              421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1825
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1945
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               601
                                                                                                                                                                                                                                                                                                                                                                                                                             361
                                                                                                                                                 ORIGIN
                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                          셤
```

EST 24-OCT-2001

linear

mRNA

700 bp

BB319509

LOCUS

ø

RESULT BB319509

```
Sciences Center(GSC), Tokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suchino-Cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9226
Fax: 81-45-503-9216
Fa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hiramoto, K., Hori, F., Ishli, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., MacQuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Tagama, M., Tagawa, A., Shinaki, T., Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa Hayashizaki,Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /clone="B230380L03"
/clone_lib="RIKEN full-length enriched, adult male corpora
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalla; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 700)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
BB319509 RIKEN full-length enriched, adult male corpora
quadrigemina Mus musculus cDNA clone B230380L03 3', mRNA sequence.
BB319509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note="Site_1: Sal1; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Rike Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Computational Analysis of Full-Length Mouse cDNAs Compared with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Laboratory for Genome Exploration Research Group, RIKEN Genomic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Human Genome Seguences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /tissue_type="corpora quadrigemina"
/dev_stage="adult"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /organism="Mus musculus"
/db_xref="taxon:10090"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /lab_host="DH10B"
                                                                                                                                                                                                                     BB319509.2 GI:16403267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    quadriqemina"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ,Y. and Hayashizaki,Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 prepare mouse tissues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /sex="male'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .700
```

```
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local
                                                                                                                                                                                                                                                                                                              source
    DEFINITION
                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNT
                               ACCESSION
VERSION
                                                                                                  REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1662
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           272
                                                                                                                                                                      TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1780
                                                                                                                                                                                        MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     332
                                                  KEYWORDS
                                                                                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE CO
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ă
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
ä
                                                                                                                                                                                                                                                                                                                                                                   1395 tttgtgcatagcccacttctgaaaaacaagggaacagtgtgtaaggaacctgtgcacatc 1454
                                                                                                                                                                                                                                                                                       1455 actgactggtaccccactctcatttcactggctgaaggacagattgatgaggacattcaa 1514
                                                                                                                                                                                                                                                                                                                               ctagatggctatgatatctgggagaccataagtgagggtcttcgctcaccccgagtagat 1574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        aaccggtggcacaatgaacggatcacctcgtcaactggcaaaagtgtatggcttttcaac 1814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1815 atcacagccgacccatatgagagggtggacctatctaacaggtatccaggaatcgtgaag 1874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1875 aageteetaeggaggeteteaeagtteaaeaaaetgeagtgeeggteaggtateeeee 1934
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           aaagaccccagaagtaaccctaggctcaatggaggggtctagggaccatggtatagagag 1994
                                                                                                                                                                                                           1335 aactggcctctcagaggtagcaaaggaacatattgggaaggagggatccgggctgtaggc 1394
                                                                                                                                                                                                                                                                                                                                                                                                                                                 479
                                                                                                                                                                                                                                                            62 TTTGCGCATAGCCCACTTTTAAAAACCAAGGACCGGGTATGTAAGGAACTTGT-CACATC 120
                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                      3 ATCTGCCCCCTCAAGGGCAGCAAAGGAACATATT-GGAAGGGGGCACCCGGGCAGTTGGC 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CCCGATTGGGACCCAACCTTGATTTCCCTGGCTGAAGGACAGATTGAT-AAACCACTCAG
                                                                                                                                                                                                                                                                                                                                          180 GTAGATGGATACGATATCTGGGAGACCATAAGCGAAGGTCTTCGTTCACCCCGAGTGGAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .,
;
                                                                                                                                                                      Length 700;
                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gaaaccaagaaaaagaagccaagcaaaaatcaggctgagaa 2035
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GAAAACAAGAAAAAGAAGTCAAACAAAACCAAGGCTAAGAA 700
                                                                                                                                                                      ..
                                                                                                                                                                                        66
                                                                                                                                                                    Score 506.6; DB 9
Pred. No. 2.3e-87;
); Mismatches 99;
                                                                                                                               u
                                                                                                                               Б
                                                                                                                               182
                                                                                                                                                                                        ;
0
                                                                                                                                                                    Query Match 22.4%;
Best Local Similarity 85.4%;
Matches 599; Conservative
                                                                                                                             179 c
                                                                                                                                ď
                                                                                                                               204
                                                                                                                              BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                              1515
                                                                                                                                                                                                                                                                                                                                                                                                            1635
                                                                                                                                                                                                                                                                                                                                                                                                                                                 1695
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1755
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1935
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             009
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                480
                                                                                                                                                                                                           ô
                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

EST 16-DEC-1996

linear

MRNA

463 bp

AA159953

RESULT 7 AA159953/c LOCUS

```
1 (bases 1 to 063)
Hiller, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chissoe, S., Dietritch, N., DuBuque, T., Favello, A., Gish, W., Hawkins
M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore
, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
Schellenberg, K., Soares, M.B., Tan, F., Thierry-Meg, J., Trevaskis, E.,
Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Warra, M.
Generation and analysis of 280,000 human expressed sequence tags
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .,
4
                                                                                                                                                                    Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catarrhini, Hominidae, Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.ilnl.gov) for further information.
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 369.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cotogicaaciggcaaaagigitaiggoiiiticaacaicacagccgacccaiaigagaggg 1839
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           tggacctatctaacaggtatccaggaatcgtgaagaagctcctacggaggctctcacagt 1899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1900 tcaacaaaactgcagtgccggtcaggtatcccccaaaagaccccagaagtaaccctaggc 1959
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ccaaggcaaaaaatggctcctgggcagcaratgggatctggaacactgcaatccag 1661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          452 CCAAGGCAAAAAATGGCTCCTGGGCAGCAGCAGCTTATGGGATCTGGAACACTGCAATCCAG 393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
    Homo sapiens cDNA clc
P15848 ARYLSULFATASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ٠.
                                                                                                                                                                                                                                                                                                                                                                                                                Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 463;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 393.8; DB 9;
Pred. No. 1.2e-65;
0; Mismatches 7;
zo75a09.sl Stratagene pancreas (#937208)
IMAGE:592696 3' similar to SW:ARSD_HUMAN
PRECURSOR ;, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1. .463
/organism="Homo sapiens"
/db_xref="GDB:4624082"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /db_xref="taxon:9606"
/clone="IMAGE:592696"
                                                            AA159953
AA159953.1 GI:1734444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         17.48;
97.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches 442; Conservative
                                                                                                                                                   Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ø
                                                                                                                                                                                                                                                                                                                                                                                                   97044478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         82
                                                                                                                               human.
                                                                                                          EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

~

```
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904 e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G., Pertea, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and Keele, J.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EST 25-APR-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /db_xref="taxon.9913"
/clone_lib="MARC 3BOV"
/tissue_type="pooled"
/lab_host="blibB"
/note="Vector: pCMV SPORT6; Site_1: Xba1; Site_2: Xho1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle Genome Res. 11 (4), 626-630 (2001)
                                                                                                                                                                                        1685 gaaattgcttacaggaaatcctggctacagcgactgggtcccccctcagtctttcagcaa 1744
       1625 ggcagcaggctatgggatctggaacactgcaatccagtcagccatcagagtgcagcactg 1684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   277595 MARC 3BOV Bos taurus cDNA 5', mRNA sequence. BF653682 GI:11918814 EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Contact: Smith FPL USDA, ARS, US Meat Animal Research Center USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: Smith@email.marc.usda.gov
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BACKWARD: GTTTTCCCAGTCACGACG
Plate: 66 row: F column: 20
Seg primer: ATTTAGGTGACACTATAG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1. .459
/organism="Bos taurus'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FORWARD: AGGAAACAGCTATGACCAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bos taurus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM
                                                                                                                                          353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1925
                                                  413
                                                                                                                                                                                          1745
                                                                                                                                                                                                                                                                                    1805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1985
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
BF653682
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                         Dp
                                                                                                δ
                                                                                                                                          g
                                                                                                                                                                                        ò
                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            414 bp mRNA linear EST 17-OCT-1996 2487a04.s1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone IMAGE:347598 3' similar to SW:ARSB_HUMAN P15848 ARYLSULFATASE B PRECURSOR; , mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 (bases 1 to 414)
Hillatri. Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, P., Trevaskis, E., Waterston, R., Willamson, A., Wohldmann, P. and Wilson, R.
The Washu Merck EST Project
Unpublished (1995)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Contact: William IX
Washington University School of Medicine
4444 Forest Parkway, Box 8501, St. Louis, MO 63108
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Fax: 314 286 1800
Fax: 314 286 1810
Fax: 31
                                                                        tcaatggaggggtctagggaccatggtatagagaggaaaccaagaaaaagaagccaagca 2019
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             э
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 368.8; DB 10; Length 414;
Pred. No. 7.6e-61;
0; Mismatches 7; Indels 3;
                                                                                                                                                           105 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          16.3%;
illarity 97.6%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                    W81455
W81455.1 GI:1392485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   105 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 406; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM
                                                                      1960
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE
JOURNAL
                                                                                                                 93
                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                               RESULT (
W81455/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMMENT
                                                                                                               셤
                                                                                                                                                             ò
                                                                                                                                                                                                      g
                                                                      ò
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3k,N.)
Hillier,L., Clask,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,T.R., Williamson,A., Wohldmann,P. and Wilson,R.
Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Juliana Brown Seq Primer: -40RP from Gibco High quality sequence stop: 482.
                                                                                                                                                                                                                                                                                                                                                 /tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/dev_stage="Adult"
/lab_nost="TOPIO"
/lab_nost="TOPIO"
/note="Organ: Pancreas; Vector: pZErO-2; Site_1: Not I;
Site_2: Xho I; Library constructed using SuperScript
Plasmid Library kit (Life Technologies). cDNA made by
oligo-dr priming. Xho I site destroyed during cloning.
Size-selected by column fractionation; average insert
size 1.59 kb. Primary library, unamplified."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        agttcaacaaaactgcagtgccggtcaggtatccccccaaagaccccagaagtaacccta 1956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2d87a04.rl Soares_fetal_heart_NbHH19W Homo sapiens CDNA clone IMAGE:347598 5' similar to SW:ARSB_HUMAN P15848 ARYLSULFATASE WB1485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1957 ggctcaatggagggtctagggaccatggtatagaggaaaccaagaaaaagaagccaa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 341.8; DB 10;
Pred. No. 1.1e-55;
0; Mismatches 2;
                                                                                                                                                                                                           1. .547
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Melton Human Is
/sex="Both"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15.1%;
99.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         W81485
W81485.1 GI:1392515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  354;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 11
W81485
LOCUS
DEFINITION
                                                                                                                                                                                                                       source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AUTHORS
                                                                                                                                                                                       FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mammalia: Butheria: Craniata; Vertebrata; Buteleostomi; Mammalia: Butheria: Primates; Catarrhin1: Hominidae; Homo.

1 (Bases I to 547)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Scearce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillart, L., Marra, M., Pape, D., Wylle, T., Marthn, J., Blistain, A., A., Glbbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T., Endocrine Pancreas Consortium
Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BG661302 547 bp mRNA linear EST 05-JUL-2001 1459e07.yl Melton Human Islets HIZ1 Homo sapiens CDNA 5', mRNA
           Library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendonosus muscle, and fetal longissimus muscle." 119 g 104 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Endocrine Parizeas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
AM 02138
Tel: 617-495-1812
Fax: 617-495-8557
                                                                                                                                                                                                                                                                                                                    ccctcagtctttcagcaacctgggaccgaaccggtggcacaatgaacggatcacctcgt 1785
                                                                                                                                                                                                                                                                                                                                                                                                                                      1486 ctgaaggacagattgatgaggacattcaactagatggctatgatatctgggagaccataa 1545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gtgagggtcttcgctcaccccgagtagatattttgcataacattgaccccatatacacca 1605
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      caactggcaaaagtgtatggcttttcaacatcacagccgacccatatgagagggtggacc 1845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1111 | 1 | 1111 | 1 | 1111 | 1 | 1111 | 1 | 1111 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                              DB 10; Length 459;
                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                              Score 361.4; DB 10;
Pred. No. 1.9e-59;
0; Mismatches 61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1846 tatctaacaggtatccaggaatcgtgaagaagctcctac 1884
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Email: dmelton@biohp.harvard.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BG661302.1 GI:13816937
                                                                                                                                                                                                        Query Match
Best Local Similarity 86.7%;
Matches 398; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens
                                                                                                  ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sequence.
BG661302
                                                                                                118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  human.
                                                                                             BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                241 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1786
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
BG661302
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMMENT
```

g

ò g

1;

Gaps

ï

Length 547; Indels

g

à g õ pp à

ò 셤 à a ò 쉽 õ

Islets

2016

9

EST 17-0CT-1996

æ

2195

2135

JOURNAL COMMENT

SOLLOS

FEATURES

```
On May 8, 1995 this sequence version replaced gi:801403.
On May 8, 1995 this sequence version replaced gi:801403.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1275
High quality sequence stops: 384 Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL: contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1275
Std Error: 0.00
Seq primer: Promaga -21m13
High quality sequence stop: 384.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bukaryotz, Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases I to 458)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.
Unpublished (1995)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lafmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1866 atcgtgaagaagtcctacggaggctctcacagttcaacaaaactgcagtgccggtcagg 1925
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1746 ctgggaccgaaccggtggcacaatgaacggatcacctcgtcaactggcaaaagtgtatgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1806 cttttcaacatcacagccgacccatatgagagggtggacctatctaacaggtatccagga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 10; Length 458;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     26; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /organism="Homo sapiens"
/db_xref="GDB:405201"
/db_xref="taxon:9606"
/clone="IMAGE:32884"
/clone_lib="Soares infant brain INIB"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 331.2; DB 1
Pred. No. 1.2e-53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /sex="female"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    14.68;
92.68;
                     GI:825405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                        Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ø
                  R43179.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 74
                                                        human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       401;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BASE COUNT
ORIGIN
                                                                        ORGANISM
                                                                                                                                                                                                               TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
     ACCESSION
                   VERSION
KEYWORDS
SOURCE
                                                                                                                         REFERENCE
                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                 COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                    ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EST 22-MAY-1995
                                                                                                  Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
This clone is available royalty-free through LLNL; contact the
TMAGE Consortium (info@inage.llnl.gov) for further information.
Insert Length: 546 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 303.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1551 ggtcttcgctcaccccgagtagatattttgcataacattgaccccatatacaccaaggca 1610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                458 bp mRNA linear EST 22-MAY-
yg18d03.s1 Soares infant brain 1NTB Homo sapiens cDNA clone
IMAGE:32854 3' similar to SP:ARSB_HUMAN P15848 ARYLSULFATASE B
PRECURSOR;, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0;
                              Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 333.2; DB 10; Length 344;
Pred. No. 5.5e-54;
0; Mismatches 6; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1851 aacaggtatccaggaatcgtgaagaagctcctacggaggct 1891
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AACAGGTATCCATGANTCGTGAAGANGCTCCTACGGGGGCT 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 74 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     б
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   14.78;
98.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 O
Unpublished (1995)
Contact: Wilson RK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match 14.7
Best Local Similarity 98.2
Matches 335; Conservative
                                                                                                                                                                                                                           1. .344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NDHL19W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 92
```

BASE COUNT ORIGIN

1611 61

ð q ò g 1731

ò

1671

å

g

1791

ò g ò

241

181

g

5

Gaps

9

1745

1805

1865

RESULT 12 R43179/c LOCUS

301

a

```
1050 actcagagagtacagcaaatcttag 1074
                                         BM461300
BM461300.1 GI:18510340
                                                                                                                                                                                                                                                                                          Unpublished (1999)
                                                                                                                                                                                                                 Mus musculus
                                                                                                                                                                                                  house mouse.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                351
                                                                                                                      LOCUS
DEFINITION
                                                    301
                                                                                                                                                                                                             ORGANISM
                                                                                                                                                                                                                                                                AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1956
                                                                                                                                                           ACCESSION
                                                                                                                                                                      VERSION
KEYWORDS
                                                                                                                                                                                                                                                      REFERENCE
                                                                                                       BM461300
                                                                                                                                                                                                                                                                                                                                                                                                                                                       FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORIGIN
                          ŏ
                                                  Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                         U46280 Human pancreatic cancer cell line Patu 8988t Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="xs254"
/clone=lb="Human pancreatic cancer cell line Patu 8988t"
/tissue_type="pancreatic cancer"
/cell_line="Patu 8988t"
1 68 c 73 g 83 t 4 others
                                                                                                                                                                                                                                                                                                                                                    (bases 1 to 325)
Gress, T.M., Muller-Pillasch, F., Geng, M., Zinmerhackl, F., Zehetner, G., Friess, H., Buchler, M., Adler, G. and Lehrach, H.
A pancreatic cancer-specific expression profile
Oncogene 13 (8), 1819-1830 (1996)
                                                                                                                                                                                                                                                                                                                    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                                                                           tatgaaaacgacaatgctgcctgggactatgacaatggcatatactccacacagatgtac 1049
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AATCGTGAAGANAGNCTACGGAGGCTCTCACAGTTCAACAAAACTGCAGTGCCGGTCAGG 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  751 ctatcataagacctacccaacccaactgtttacctctggacaatgccaccctacctcaga 810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 AACTGAAGGAGGTTGGNTATTCAACGCATATGGTCGGAAAATGGCACTTGGGTTTTTACA 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Internal Medicine I. University of Ulm
Robert-Koch-Strasse 8, Ulm, Baden-Wuerttemberg 89081, Germany 7e1: 0731/502-4311
Exa: 0731/502-4302
Email: friederike.mueller-pillasch@medizin.uni-ulm.de.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 CTATCATAAGACCTACCCAACCCAACTGTTTACCTCTGGACAATGCCACCCTACCTCAGA 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 307.4; DB 10; Length 325;
Pred. No. 5e-49;
0; Mismatches 5; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                         Contact: Friederike Mueller-Pillasch
                                                                                                                                                                                                                                                   cDNA clone xs254, mRNA sequence.
U46280
U46280.1 GI:1236118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13.6%;
ilarity 98.2%;
Conservative (
                                                                                                                              2046 aaaagcaaaaaa 2058
                                                                                                                                                        AAAAGCAAAAAA 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ಹ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97
                                                                                                                                                                                                                                                                                                         human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Sir
Matches 319;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    source
                                                                                                                                                                                                                                                              ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BASE COUNT
ORIGIN
                                                                                                                                                                                                13
                                                                                                                                                                                                                                       DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                TITLE
JOURNAL
MEDLINE
COMMENT
194
                          1926
                                                                                                                                                        14
                                                                                                                                                                                                                                                                                                                                                         REFERENCE
AUTHORS
                                                                             1986
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        871
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                RESULT
U46280
                                                 q
                                                                                                    q
g
                        ò
                                                                             à
                                                                                                                              ò
                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
BM461300 1157 bp mRNA linear EST 05-FEB-2002 AGENCOURT_6419791 NCI_CGAP_Ov44 Mus musculus cDNA clone IMAGES:5504049 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1. .1157
/ Organism="Mus musculus"
/ db_xref="taxon:10090"
/ clone="lib="NCI_CGAP_Ov4"
/ lab_host="nHIOB (TI phage-resistant)"
/ lab_host="nHIOB (TI phage-resistant)"
/ note="Organ: ovary, PMSG-treated; Vector:
pCMV-SPORT6.ccdb: Site=1: EcoRv; Site=2: NotI; Cloned
unidirectionally. Primer: Oligo dT. Average insert size
2. 2 kb. Library constructed by Life Technologies. Note:
this is a NIGGAP Library."
1 a 260 c 251 g 294 t. 1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Aaron Hauch
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12144 row: m column: 10
High quality sequence stop: 738.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Euteleostomi;
Murinae; Mus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1955
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1776 atcacctcgtcaactggcaaaagtgtatggcttttcaacatcacagccgacccatatgag 1835
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1716 gactgggtccccctcagtctttcagcaacctgggaccgaaccggtggcacaatgaacgg 1775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostor
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; I
I (bases 1 to 1157)
NIH-WGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
241 TATGAAAACGACAATGCTGNCTGGGACTATGACAATGGNATATNCTCCACACAGATGTAC 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21 GACTGGGTCCCCCCTCAGGCTTTC-GCAATCTGGGCCCAAAACCGGTGGCACAAATGAAAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                cagttcaacaaaactgcagtgccggtcaggtatcccccaaagaccccagaagtaacct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       200 CAATTCAACAAGACTGCCGTGCCTGTCAGGTACCCCCCGAAGGATCCCAGAAGCAACCCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 1157;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0; Mismatches 114;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 12.6%; Score 284.6; DB 1
Best Local Similarity 75.1%; Pred. No. 8.1e-45;
Matches 402; Conservative 0; Mismatches 114
```

```
Search completed: July 23, 2002, 15:52:09
Job time: 3130 sec
716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QQ
                                                   g
                                                                                                                           δλ
                                                                                                                                                                                         g
                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                              δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BE553565 628 bp mRNA linear EST 15-AUG-2000 ur47a08.yl NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:3153398 5' similar to SW:ARSB_FELCA P33727 ARYLSULFATASE B PRECURSOR;, mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /clone_lib="WCL_CGAP_Mam2"
/tissue_type="tumor, blopsy sample"
/tissue_type="tumor, blopsy sample"
/dev_stage="5 months"
/lab_host="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blost="blo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus I Loasa; 1 to 628)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rémail.nh.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
                              2071
                                                                                                                                                          cagaaagcagtctcaggttcaacttgccattcaggtgttacttgtgggataagcacaaata 2131
                                                             380 TIGGCAGCICACICCICAAIAAAAIGCCACCCAAGIGIIGCCACIGGCIAAAIGCAAAAC 439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 440 TITCIGCCTIGGITAAACTTAGA-------CIGCCTGGITTCTAGGIAATTCA 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  596 gggttaccacggatctgagattaaaacacctactctgacaagctcgctgccgaaggagt 655
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 GGGATACCATGGCTCGGATATCGAGACCCCAACGCTGGACCGGCTAGCAGCTGAGGGCGT 60
                                                                                                                                                                                                                                                                                      2132 tttcctgtttggttaaactttaatcagttcttatctttcatctgtttcctaggtaaacca
                              agcaaaaatcaggctgagaaaaagcaaaagaaaagcaaaa----aaagaagaagaaacag
                                                                                                                                                                                                                                                                                                                                                                                                                      2192 gcaaatttggctcgataatatcgctggcctaagcgtcaggcttgttttcatgctg 2246
                                                                                                                                                                                                                                                                                                                                                                                                                                                            486 AAGAAATTGGCTCAGAAATACCTCGGCCCACGCGCTTGTTTTCACATTG 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 628;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11.6%; Score 262; DB 10; Length 628 64.6%; Pred. No. 2.1e-40; ive 0; Mismatches 220; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1. .628
/organism="Mus musculus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              quality sequence stop: 390.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /strain~"FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:3153398"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               primer: -40RP from Gibco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BE553565.1 GI:9817982
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Tumor Gene Index
Unpublished (1997)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 11.6
Best Local Similarity 64.6
Matches 407; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MGI:1056154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            house mouse
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sednence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 15
BE553565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOURCE
ORGANISM
                           2016
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
ORIGIN
                                                                                                                                                       2072
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMMENT
                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                        a
                           ð
                                                                                      g
                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
```

```
1015
                                                                                                                                                                                                                    1135
                                                                                                                                                                                1075
                                                                                                                                                                                                                                                        1136 actgcaagctcctggcaggtatttcgaacactaccgatccattatcaacataaacaggag 1195
                                                                                                                                                                                                 477
                                                                                       300
                                                                                                         955
                                                                     895
                                                                                                                                                                                                                                                                          597
                                                                                                       121 CGGCAGGTACCAGATCCATACAGGATTGCAGCACTCCATTATCCGCCCAGGGCAGCCCAA
                                            tggaaagtatcagatacacaccggacttcaacattctatcataagacctacccaacccaa
                                  776 ctgtttacctctggacaatgccaccctacctcagaaactgaaggaggttggatattcaac
                                                                                                                                                                                                                                                                                             1196 gagatatgctgccatgctttcctgcttaga 1225
```

.

. ( . ( BBS20951 BBS20951 BB119509 BB11509 BE149009 CMO-HT024 R20122 y918d03.r1 BE655682 277595 MA BE047859 £243c06.y BG975362 602843009 BT693994 H0A40-1-E BG899940 H0A40-1-E ALO22653 87827b33 AVS96948 AVS96948

BG899940

A1641294 fc13b07.y A1497032 fb59c12.y A1877708 fc50g02.y AV617615 AV617615

BF537953 602049394 BI328079 602979220 BI697980 603344126

AW429593 67921 MAR AV615802 AV615802 BM495947 IpcGBr5\_2

BM280523 ki01h08.y AV595356 AV595356 AV595699 AV595699

BI109023 602896840 BI689140 603315947 AIS45901 fb76h10.y BI684551 603306926

ALIGNMENTS

ouo.

Run

```
BG680752 817 bp mRNA linear EST 01-MAY-2001 602628445F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4753259 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /tissue_type="squamous cell carcinoma"
//tab_host="bildom (TT phage=resistant)"
//hab_host="bridom (TT phage=resistant)"
//hote="Organ: skin; Vector: pCWV-SPORT6; Site_1: Not1;
Site_2: Sal1; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 817)
                                                                                                                                                                                                                                                                                                                                                                                                                         l (bases 1 to 81/)
NIH-WGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
.... ^~~~manding by: Incyte Genomics, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can if cound through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAMJ0612 row: j column: 12
High quality sequence stop: 789.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4753259"
/clone_lib="NCI_CGAP_Skn4"
                                                                                                  BF537953
BI328079
BI697980
                                                                                                                             AW429593
AV615802
BM495947
                                                                                                                                                        AI641294
AI497032
AI877708
AV617615
BM280523
AV595356
AV595356
                                                                                                                                                                                                                         BI109023
BI689140
        BB319509
BE149009
R20122
BF653682
                                             BE047859
BG975362
                                                                                          AV596948
                                                                                                                                                                                                                                          AI545901
                                                                                                                                                                                                                                                                                                                                                                BG680752.1 GI:13912149
1. .817
                                                                                                                                                                                                                                                                                                                                                mRNA sequence.
BG680752
                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens
human.
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                  AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                            RESULT
BG680752
                                                                                                                                                                                                                                                                                                                                                         ACCESSION
                                                                                                                                                                                                                                                                                                                                                                  VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
υ
                                                                       (without alignments)
12784.552 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BE307503 601893984
BG772761 602720927
WB1485 2687a04.r1
BG661302 1a529e07.y
BF979627 602287919
BC194726 RST13892
BC194726 RST13892
BC194726 RST13892
BC194726 RST13892
BC194726 RST13892
BC194726 RST13892
BC1979627 601881811
BK6280 HSU46280 HU
WB1455 2618044.s1
RK3179 y918403.s1
AA159953 2075a09.s
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BG680752 602628445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BG167651 602342825
AZ815176 2M0083G13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AW474222 xs20h05.x
                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                          2266
1 cacgcgtccgcccacgcgtc......tgccacctggtgccgaattc 2266
                                                            July 23, 2002, 15:08:04; Search time 2392.27 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Description
                                                                                                                                                                                                      27472414
       GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                 13736207 segs, 6748477542 residues
                                                                                                                                                                                                    Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BF979627
BG194726
BF218442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BF307503
BG772761
W81485
BG661302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AA159952
BG167651
AZ815176
                                            nucleic search, using sw model
                                                                                                                                                                                                                                                    Post-processing: Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BG680752
AW474222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AA159953
BF696654
                                                                                                                                      OLIGO_NUC
Gapop 60.0 , Gapext 60.0
                                                                                                                                                                                                                                                                                                                                                                                                     em_gss_inv:*
em_gss_pln:*
em_gss_vrt:*
                                                                                                                                                                                                                       seq length: 0
seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                              em_gss_hum:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100000
                                                                                                  US-09-495-823-8
                                                                                                                                                                                                                                                                              em_estba:*
em_esthum:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB
                                                                                                                                                                                                                                                                                                 em_estin:*
em_estwu:*
em_estvo:*
em_estro:*
em_htc:*
gb_est1:*
gb_htc:*
gb_hts:*
gb_hts:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                683
344
1032
252
4124
4114
458
8463
625
625
                                                                                                                                                                                                                                                                       EST:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Perfect score:
Sequence:
                                                                                                                                        Scoring table:
                                                                                                                                                                                   Word size :
                                                                                                                                                                                                                        Minimum DB
Maximum DB
                                            OM nucleic
                                                                                                                                                                                                                                                                       Database :
                                                                                                                                                                  Searched:
```

259

BASE COUNT

000

õ

O O Center n can be

Wed

ORIGIN

1428

69

1368

ò 8 ò 8 ò g ò g ò g ò q ò q ò q å g ò g ð

1488 129 1548 1608

249 1668 309 1728 369

189

```
BF307503 683 bp mRNA linear EST 21-NOV-2000 601893984F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4139565 5',
                                                                                                                                                                     /organism="Homo sapiens"
/db_xref="taxon:9606"
/db_xref="taxon:9606"
/clone="InAGE:2770233"
/clone=lib="NcICGAP_Ut2"
/tissue_type="moderately-differentiated endometrial
adenocarcinoma, 3 pooled tumors"
/lab_host="OHIOB"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: Sal1;
Site_2: Not1; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.85 kb. Life Technologies catalog #:
                                                 Sequencing Cente
information can
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        gaaateetggetacagegaetgggteeeceteagtettteageaaeetgggaeegaaee 1750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cagccgacccatatgagaggtggacctatctaacaggtatccaggaatcgtgaagaagc 1878
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      tgcataacattgaccccatatacaccaaggcaaaaatggctcctgggcagcagctatg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 413 GGATCTGGAACACTGCAATCCAGTCAGCCATCAGAGTGCAGCACGGAAATTGCTTACAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       233 CAGCCGACCCATATGAGAGGGTGGACCTATCTAACAGGTATCCAGGAATCGTGAAGAAGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                473 TGCATAACATTGACCCCATATACACCAAGGCAAAAAATGGCTCCTGGGCAGCAGCTATG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ggatctggaacactgcaatccagtcagccatcagagtgcagcactggaaattgcttacag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ggtggcacaatgaacggatcacctcgtcaactggcaaaagtgtatggcttttcaacatca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
Numert-Buck, M.D., Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencially by: Washington University Genome S Clone distribution: NCI-CGAP clone distribution i
                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                           found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Seq primer: 400UP from Gibco
High quality sequence stop: 435.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                Length
                                                                                                                                                                                                                                                                                                                                                                                                                                           1;
                                                                                                                                                                                                                                                                                                                                                                                                           Score 465; DB 9;
Pred. No. 1e-180;
0; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                accccagaagtaaccctaggctcaatggagggtct 1974
                                                                                                                                                                                                                                                                                                                                                  ų
                                                                                                                                                                                                                                                                                                                                                  195
                                                                                                                                                                                                                                                                                                                                               153 g
                                                                                                                                                                                                                                                                                                                                                                                                               20.5%;
99.8%;
                                                                                                                                                                                                                                                                                                                                   11539-012"
                                                                                                                                                                                                                                                                                                                                               151 c
                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 20.5
Best Local Similarity 99.8
Matches 515; Conservative
                                                                                                                                                         1. .615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           mRNA sequence.
BF307503
                                                                                                                                                                                                                                                                                                                                                  116
                                                                                                                                                             source
                                                                                                                                                                                                                                                                                                                                                 COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1639
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1879
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BF307503
                                                                                                                                            FEATURES
                                                                                                                                                                                                                                                                                                                                                 BASE CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AW474222 615 bp mRNA linear EST 24-FEB-2000 xs20h05.x1 NCI_CGAP_Ut2 Homo sapiens cDNA clone IMAGE:2770233 3' similar to SW:ARSB_HUMAN P15848 ARYLSULFATASE B PRECURSOR;, mRNA
                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae, Homo.
1 (bases 1 to 615)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ď
                                                                                                                                                                                                                                                                                                                                                              1787
                                                                                                                                                                          1487
                                                                                                                                                                                                                                                                                                   1607
                                                                                                                                                                                                                                                                                                                                                                                                                            1727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     actggcaaaagtgtatggcttttcaacatcacagccgacccatatgagaggtggaccta 1847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   totaacaggtatccaggaatcgtgaagaagctcctacggaggctctcacagttcaacaaa 1907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              actgcagtgccggtcaggtatccccccaaagaccccagaagtaaccctaggctcaatgga 1967
                                                                                                                                                                                                                                      gaaggacagattgatgaggacattcaactagatggctatgatatctgggagaccataagt 1547
                                                                                                                                                                                                                                                                     188
                                                                                                                                                                                                                                                                                                                                248
                                                                                                                                                                                                                                                                                                                                                                                                                                                           368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael
                                                                                                                                                                                                       128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             608
                                                                                                                                         9
                                                                              Gaps
                                                                                                           GAAGGACAGATTGATGAGGACATTCAACTAGATGGCTATGATATCTGGGAGACCATAAGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
0
                                                                              Indels
                                                 Length
                                               Score 505; DB 10;
Pred. No. 3.1e-197;
                                                                            0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AW474222.1 GI:7044328
                                             22.3%;
99.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Tumor Gene Index
Unpublished (1997)
                                             Query Match 22.3
Best Local Similarity 99.7
Matches 605; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ggggtct 1974
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sequence.
AW474222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GGGGTCT
```

1788

429

1908

489

1848

ö

Gaps

1638

1698

354

1818

174

LOCUS

ACCESSION VERSION KEYWORDS

AW474222/C

RESULT

609

1968

SOURCE ORGANISM

AUTHORS TITLE

JOURNAL

COMMENT

REFERENCE

m

```
The following following the following the following the following following the following the following the following following the following the following followin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       W81485 344 bp mRNA linear EST 17-OCT-1996 zd87a04.rl Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (gtcgag
                                                        BG772761 T14_MGC_97 Homo sapiens cDNA clone IMAGE:4837638 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library." 235 c 185 g 146 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note-"Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcg_1); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTVN-3' size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the
                                                                                                                                                                                                                                                           Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ttttttttttccttggtggaagctgctctagggaagggggggaggaggaggaggaagtgaaa 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 237 TITITITITITICCTIGGIGGAAGCTGCTCTAGGGAGGGGGGGGGAGGAGGAAGTGAAA 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 735;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .,
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9.4%; Score 212; DB 10;
100.0%; Pred. No. 1.2e-76;
Live 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    224 caggtggagagtcttcaccgtctgttgggtgc 255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /organism="Homo sapiens"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /db_xref="taxon:9606"
/clone="IMAGE:4837638"
/clone_lib="NIH_MGC_97"
/lab_host="DH108"
                                                                                                             mRNA sequence.
BG772761
BG772761.1 GI:14083414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 212; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    169 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
  RESULT 4
BG772761/c
                                                                                  DEFINITION
                                                                                                                                                                                                                                                              ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          164
                                                                                                                                                                                                                                                                                                                                             REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               44
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            57
                                                                                                                                             ACCESSION
                                                                                                                                                                        VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            W81485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCALLY...

1. ..683

/Organism="Homo sapiens"
/Organism="Homo sapiens"
/Clone="IMAGE:4139565"
/Clone="IMAGE:4139565"
/Clone="IMAGE:4139565"
/Clone="IMAGE:4139565"
/Clone="Corgan: muscle; Vector: pOTB7; Site_1: ECORI; Site_2: XhOI; CDNA made by Oligo-dr priming.
Directionally cloned into ECORI/XhOI sites using the following 5' adaptor: GGCAGGAGGG(G). Size-selected >500pp for average insert size 1.8kb. Library constructed by Ling Homos in the laboratory of Gerald M. Rubin (University of Callifornia, Berkeley) using ZAP-CDNA synthesis kit (Stratagane) and Superscript II RI (Life Technologies)."
                                                                                                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; E 1 (bases 1 to 683)

1 (bases 1 to 683)

2 NIH-MCC http://mgc.nci.nih.gov/.

3 NIH-MCC http://mgc.nci.nih.gov/.

4 Unpublished (1999)

5 Contact: Robert Strausberg, Ph.D.

5 Email: cgapbs-remail.nih.gov

7 Tissue Procurement: ATCC

6 CDNA Library Preparation: Ling Hong/Rubin Laboratory

6 CDNA Library Preparation: Ling Hong/Rubin Laboratory

7 CDNA Library Preparation: Ling Hong/Rubin Laboratory

8 CDNA Library Preparation: Ling Genomics, Inc.

8 CLONG ALIBRITION MCC CLONG GISTRIBULION INFORMATION HONG HOUGH LIP LINA G.E. CONSORTIUM/LINI at: image.llnl.gov

9 Plate: LLCM1050 row: g column: 22

10 High quality sequence stop; 679.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     122 GAGTTAAACTGGAGAACTACTATGTCCAGCCTATTTGCACCATCCAGGAGTCAGTTTA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gagttaaactggagaactactatgtccagcctatttgcacaccatccaggagtcagttta 711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      712 ttactggaaagtatcagatacacaccggacttcaacattctatcataagacctaccaac 771
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ccaactgittaccictggacaatgccaccctaccicagaaactgaaggaggitggatati 831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      592 atgigggitaccacggaictgagaitaaaacacciactcttgacaagcicgctgccgaag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 683;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16.9%; Score 382; DB 10; Length 68
100.0%; Pred. No. 1.4e-146;
Live 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BF307503.1 GI:11254633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches 382; Conservative
                                                                                     Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                  ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         652
                                                                                                                                                               REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    772
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                  COMMENT
                                                     SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

ö

```
size 1.59 kb. Prima
124 c 120 g
                                                                                                                                                                                                                                   BG661302.1 GI:13816937
                                                              1851 aacaggtatcca 1862
                                                                              301 AACAGGTATCCA 312
                                                                                                                                                                                                         sedneuce
                                                                                                                                                                                                                      BG661302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         169
                                                                                                                                                                                                                                                                   human.
                                                                                                                                            9
                                                                                                                                                                                    DEFINITION
                                                                                                                                                                                                                                                                                 ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
JOURNAL
                                                                                                                                                                                                                      ACCESSION
                                                                                                                                                                                                                                                                                                                             REFERENCE
                                                                                                                                            RESULT
BG661302
                                                                                                                                                                                                                                    VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMMENT
                                                                                                                                                                         LOCUS
                              g
                                                              ŏ
                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mammalia Eutimeria; Filmderes; Cacatiniii, Nominiatore, Nominia.

E 1 (bases 1 to 344)

S Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman

N., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,

Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston

Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston

Rifkin, L., Wohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston

Rifkin, L., Wohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston

R., Williamson, A., Wohldman, P. and Wilson, R.

Rashington University School of Medicine

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1810

Email: est@atson.wustl.edu

This clone is available royalty-free through LLML; contact the

This clone is available royalty-free through LLML; contact the

This clone is available royalty-free through LLML; contact the

This clone is available royalty-free through LLML; schartion.

Insert Length: 546 Std Error: 0.00

Seq primer: mob.REGA+ET

High quality sequence stop: 303.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
0
                                                                                          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Д
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      agagtgcagcactggaaattgcttacaggaaatcctggctacagcgactgggtccccct 1730
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IMAGE: 347598 5' similar to SW: ARSB_HUMAN P15848 ARYLSULFATASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121 AGAGTGCAGCACTGGAAATTGCTTACAGGAAATCCTGGCTACAGCGACTGGGTCCCCCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 344;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 210; DB 10;
Pred. No. 1.1e-75;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                74 t
                                                                                                                                                                                                                                                                                                                                                                                                                                    1. .344
/organism="Homo sapiens"
/db_xref="GDB:1272973"
/db_xref="taxon:9606"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              82 g
                PRECURSOR ;, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9.38;
                              W81485
W81485.1 GI:1392515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 99.4
Matches 310; Conservative
                                                                              human.
                                                                                                                                                                                                                                                                                                                                                                                                                                       source
                                                                          SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BASE COUNT
ORIGIN
                                                                                                                                                                                                                 TITLE
JOURNAL
COMMENT
                                                                                                                                       REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1671
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1731
                             ACCESSION
                                                             KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURES
                                              VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalla; Eutherla; Primates; Catarrhini; Hominidae; Homo.

1 (Dases I to 547)

1 (Dases I to 547)

2 Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Scearce, M., Baestelli, J., Gradwohl, G., Clifton, S., Hiller, L., Marra, M., Pape, D., Wylle, T., Marth, J., Blistain, A., Chalsing, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T., Jackson, Y. and Bowers, Y.

Endocrine Pancreas Consortium

Lontact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, M. M. 01318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
BG661302 547 bp mRNA linear EST 05-JUL-2001 ia59e07.yl Melton Human Islets HIZ1 Homo sapiens cDNA 5', mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Tel: 617-495-1812
Fax: 617-495-857
Fax: 617-495-857
Fax: 617-495-857
Email: dmelton@blohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on obtaining a clone please contact: Juliana Brown
(brownefass.harvard.edu)
Seq primer: -40RP from Gibco
High quality sequence stop: 482.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note-"Organ: Pancreas; Vector: p2ErO-2; Site_1: Not I; Site_2: Xho I; Library constructed using SuperScript Plasmid Library constructed using SuperScript Plasmid Library Lit (Life Technologies). CDNA made by oligo-dr priming. Xho I site destroyed during cloning. Size-selected by column fractionation; average insert size I:59 kb. Primary library, unamplified."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2112 cttgtggataagcacaaatatttcctgtttggttaaactttaatcagttcttatcta 2171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 8.9%; Score 201; DB 10; Length 547; Best Local Similarity 100.0%; Pred. No. 4.5e-72; Matches 201; Conservative 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /clone_lib-"Melton Human Islets HIZ1"
/sex-"Both"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="TOP10"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /organism="Homo sapiens"
/db_xref="taxon:9606"
```

S

셤

ð

a

ò

```
BF218442 552 bp mRNA linear EST 06-NOV-2000 601881811F1 NIH_MGC_57 Homo sapiens cDNA clone IMAGE:4094211 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Talliant 167, Libraria, Filmaces, Catainnii, nominidae, nomin.

Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J.,
Lerner, L., Costanzo, D., McElligott, K., Boozer, S., Mays, R., Smith,
E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K., Offenbacher, J., Danzig, J. and Ducar, M.

E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K., Offenbacher, J., Danzig, J. and Perotein expression libraries using random activation of gene expression

Nat. Biotechnol. 19 (5), 440-445 (2001)

Contact: Scott J. Cain
                                                                                                                                                                                                                                                      21-APR-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
/note="See 'Creation of Genome-wide Protein Expression
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                      BG194726 21-APR-2
RST13892 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
BG194726
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   180 AATGGCACTTGGGTTTTTACAGAAAGAATGCATGCCCACCAGAAGAGGGATTTGATACCT 239
     136 GICCCATCCATTAAGCCATCACTTCTGGAAGATTAAAGTTGTCGGACATGGTGACAGCTG 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  850 aatggcacttgggtttttacagaaagaatgcatgcccaccagaagaggatttgatacct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 10; Length 274;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3201 Carnegie Ave, Cleveland, OH 44115, USA Tel: 216 431 9900
Fax: 216 31 9596
Email: scain@athersys.com
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6.8%; Score 155; DB 10;
100.0%; Pred. No. 5e-53;
ative 0; Mismatches (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /organism="Homo sapiens"
/db_xref="taxon:9606"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             High quality sequence stop: 2
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                       BG194726.1 GI:13716413
                                                          76 AGAGGAGGAGGATTTCTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity 100.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Athersys, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11
                                                                                                                                                                                                                                                                                                                                                                                                    human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               155;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local S
Matches 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DEFINITION
                                                                                                                                                                                                                                                                                DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
BF218442
LOCUS
                                                                                                                                                                                                                            BG194726
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE
                                                                                                                                                                                                                                                                                                                ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOURNAL
                                                                                                                                                                                                                                                                                                                                                                     KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                             VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORIGIN
                                                                                                                                                                                               RESULT
                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
     g
                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1032)

1 (bases 1 to 1032)

1 National Institutes of Health, Mammalian Gene Collection (MGC)

1 Inpublished (1999)

1 Contact: Robert Strausberg, Ph.D.

2 Contact: Robert Strausberg, Ph.D.

3 Contact: Robert Strausberg, Ph.D.

4 Contact: Robert Strausberg, Ph.D.

5 Contact: Robert Strausberg, Ph.D.

6 Contact: Robert Strausberg, Ph.D.

7 Contact: Robert Strausberg, Ph.D.

7 Contact: Robert Strausberg, Ph.D.

8 Contact: Robert Ph.D.

9 Cont
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /note="Organ: testis; Vector: pBluescriptR (modified bpluescript KS+); Site_1: BamHi; Site_2: Sall *Nol (gtcgap b); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (WIMHYRIGKI, National Institutes of Health). Note: this is a NIH_MGC Library. 296 c 278 g 197 t lothers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                  BF979627 1032 bp mRNA linear EST 23-JAN-2001 602287919F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4373740 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         217 CTTGTGGATAAGCACAAATATTTCCTGTTTGGTTAAACTTTAATCAGTTCTTATCTTTCA 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gtcccatccattaagccatcacttctggaagattaaagttgtcggacatggtgacagctg 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 1032;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 200; DB 10;
Pred. No. 9.1e-72;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity 100.0%; Pred. No. 9.1r
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /organism-"Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4373740"
/clone_lib="NHLMGC_97"
/lab_host="DH10B"
                                                                                                                                                                     cttgttttcatgctgtgccac 2252
                                                                                                                                                                                               337 CTTGTTTCATGCTGTGCCAC 357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BF979627.1 GI:12346751
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 100. Matches 200; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1. .1032
                                                                                                                                                                                                                                                                                                                                                                                                                            mRNA sequence.
BF979627
                                                                                                                                                                                                                                                                                                                                                               BF979627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                     RESULT 7
BF979627/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                source
                                                                                                                                                                                                                                                                                                                                                                  LOCUS
DEFINITION
                                                                                                                                                                  2232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       142
                                                                                                                                                                                                                                                                                                                                                                                                                                                ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORIGIN
```

ö

ð 셤 à a ò Q

```
Mammalial is further as, continued as a mammalial is further as a manner as a mann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       414 bp mRNA linear EST 17-OCT-1996 ad87a04.sl Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone IMAGE:347598 3' similar to SW:ARSE_HUMAN P15848 ARYLSULFATASE B PRECURSOR; , mRNA sequence.
                                                                                                                                                                                                                                                                           /organism="Homo sapiens"
/db_xref="taxon:9606"
/db_xref="taxon:9606"
/clone=lib="Human pancreatic cancer cell line Patu 8988t"
/tissue_type="pancreatic cancer"
/cell_line="Patu 8988t"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     811 aactgaaggaggttggatattcaacgcatatggtcggaaaatggcacttgggtttttaca 870
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A pancreatic cancer-specific expression profile Oncogene 13 (8), 1819-1830 (1996) 9705070 Contact: Friederike Mueller-Pillasch Internal Medicine I, University of Ulm Robert-Koch-Strasse 8, Ulm, Baden-Wuerttemberg 89081, Germany Tel: 0731/502-4302
                                                                                                                                                                                                         Email: friederike.mueller-pillasch@medizin.uni-ulm.de
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 5.1%; Score 115; DB 10; Length 325; Best Local Similarity 99.4%; Pred. No. 1.3e-36; Matches 165; Conservative 0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        871 gaaaagaatgcatgcccaccagaagaggatttgataccttttttgg 916
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 83 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /organism="Homo sapiens"
/db_xref="GDB:1272973"
                                                                                                                                                                                                                                                                                                                                                                                                                             73 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                W81455.1 GI:1392485
                                                                                                                                                                                                                                                                                                                                                                                                                           68 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     human.
          TITLE
JOURNAL
MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                             BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 11
W81455/c
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FEATURES
                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                         325 bp mRNA linear EST 13-MAR-1997
HSU46280 Human pancreatic cancer cell line Patu 8988t Homo sapiens
CDNA clone xs254, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostoml;
Mammalia; Eutherla; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 325)
Gress, T.M., Muller-Pillasch, F., Geng, M., Zimmerhackl, F., Zehetner, G., Friess, H., Buchler, M., Adler, G. and Lehrach, H.
                                                                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 552)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 552;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 5.3%; Score 119; DB 10; Best Local Similarity 100.0%; Pred. No. 2.5e-38; Matches 119; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Laboratories (Palo Alto, CA)."
119 c 105 g 186 t
          BF218442
BF218442.1 GI:11112028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              U46280.1 GI:1236118
                                                                                                            Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  142
                                                                                     human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 10
U46280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
ORIGIN
                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM
                                                                                                                                                                                                                             TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
AUTHORS
            ACCESSION
                                   VERSION
KEYWORDS
                                                                                                                                                                                                         AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                REFERENCE
```

g ò g

ö

Gaps ö

FEATURES

COMMENT

SOURCE

```
2075a09.S1 Stratagene pancreas (#937208) Homo sapiens cDNA clone IMAGE:92696 3' similar to SW:ARSB_HUMAN P15848 ARYLSULFATASE B PRECURSOR ; mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 463)
1 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
TMAGE consortium (info@image.llnl.gov) for further information.
Seq primer: -40Ml3 #vd. from Amersham
High quality sequence stop: 369.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1881 ctacggaggctctcacagttcaacaaaactgcagtgccggtcaggtatcccccaaagac 1940
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                179 CTACGGAGGCTCTCACAGTTCAACAAAACTGCAGTGCGGGTCAGGTATCCCCCCAAAGAC 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ó:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /db_xref="taxon:9606"
/clone="twk8E:592696"
/clone="twk8E:5tratqene pancreas (#937208)"
/lab_host="SOLR cells (kanamycin resistant)"
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 458;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     4.1%; Score 94; DB 10; Length 45
100.0%; Pred. No. 5.2e-28;
Live 0; Mismatches 0; Indels
                                               /clone_lib="Soares infant brain 1NIB"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /organism="Homo sapiens"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /db_xref="GDB:4624082"
/db_xref="taxon:9606"
                           /clone="IMAGE:32854"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
                                                                     /sex="female"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AA159953.1 GI:1734444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  97044478
Contact: Wilson RK
                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match
Best Local Similarity 100.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                .463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens
                                                                                                                                                                                                                                                                                                                                                            74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          human.
                                                                                                                                                                                                                                                                                                                                                         BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AA159953/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                           EST 22-MAY-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1275
High quality sequence stops: 384 Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1275 Std Error: 0.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  458 bp mRNA linear EST 22-MAY. yg18d03.s1 Soares infant brain lNLB Homo sapiens cDNA clone IMAGE:32854 3' similar to SP:ARSB_HUMAN P15848 ARYLSULFATASE BPRECURSOR;, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 414;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                   /clone="IMAGE:347598"
/clone_lib="Soares_fetal_heart_NbHH19W"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 10; Lo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                130 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 4.3%; Score 97; DB Best Local Similarity 100.0%; Pred. No. 3.1 Matches 97; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /organism="Homo sapiens"
/db_xref="GDB:405201"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Seq primer: Promega -21ml3
High quality sequence stop: 384.
Location/Qualifiers
/db xref="taxon:9606"
                                                                                                                                                                                                                                                                                                                                                                                                                            105 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       R43179.1 GI:825405
                                                                                                                                                                                                                                                                                                                                                                                                                            105 c
                                                                                                                                                                                                                                                                                                                                                                                                      NbHL19W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                            74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 12
R43179/c
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                          BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMMENT
```

g ó g

ó

ö

œ

```
82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Query Match
                          Best Local
                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JOURNAL
MEDLINE
                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                  VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AUTHORS
                                                                                                                                                                                                                                                                                                                               AA159952
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMMENT
                                                                                                                                                                                         ò
                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                               δ
                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1. .849

/organism="Homo sapiens"
/db_xref="Homo sapiens"
/db_xref="Homo sapiens"
/db_xref="Homo sapiens"
/clone=lib="NIH_MGC_56"
/clone=lib="NIH_MGC_56"
/tissue_type="primitive neuroectoderm"
/lab_host="DHIOB (T1 phage-resistant)"
/note="Organ: brain: Vector: pDNR-LIB (Clontech); Site_1:
Sfil (gpccgcccggcc); Site_2: Sfil (gpccattatggcc);
Sfil (gpccatcagcc); Site_2: Sfil (gpccattatggcc);
adaptors were used in cloning as follows: 5'
adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ArcTAGAGACCGACATGATGGCC-3' and 3' adaptor
sequence: 5'-ArCTAGAGACGCGACGACACATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, Or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 collonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
E I (bases 1 to 849)

E I (bases 1 to 849)

I (bases 1 to 849)

L Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rémail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: CLONETECH Laboratories, Inc.

CLONE distribution: MGC Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

plate: LLCM1110 rown d column: 16

High quality sequence stop: 579.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BF696654 12-DEC-2000 mRNA linear EST 22-DEC-2000 602124758F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4281567 5',
                                                                                                                                                                                                                                                                                      ;;
0
  /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                 1881 ctacggaggeteteacagtteaacaaaactgeagtgeeggteaggtateeeceeaaagae 1940
                                                                                                                                                                                                                                                                                                                                                        112 CTACGGAGGCTCTCACAGTTCAACAAACTGCAGTGCCGGTCAGGTATCCCCCCAAAGAC 113
                                                                                                                                                                                                                                                                                      0; Gaps
                                                                                                                                                                                                                                       Length 463;
                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                     Score 94; DB 9; Le
Pred. No. 5.1e-28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Laboratories (Palo Alto, CA)."
175 c 180 g 257 t
                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 4.1%; Score 94; DB Best Local Similarity 100.0%; Pred. No. 5.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BF696654.1 GI:11982062
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      mRNA sequence.
BF696654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens
                                                                                                                                             ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 1
BF696654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURES
                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
```

```
Hillfart, L. Tennon, G., Becker, M., Bonaldo, M. F., Chiapelli, B., Hillart, L., Lennon, G., Becker, M., Bonaldo, M. F., Chiapelli, B., Tohiso, S., Dietrich, N., DuBuque, T., Favello, A., Gish, M., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M. B., Tan, F., Thierry-Meg, J., Trevaskis, E., Underwood, K., Wohldman, P., Waterston, R., Wilson, R. and Marra, M. Generation and analysis of 280,000 human expressed sequence tags 97044478
                                                                                                                                                                                                                                                                                                                                                                 AA159952 407 bp mRNA linear EST 16-DEC-1996 2075809.rl Stratagene pancreas (#937208) Homo sapiens cDNA clone IMAGE:592696 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Contact: Wilson RK
Washington University School of Medicine
Washington University School of Medicine
4444 Forest Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1810
Fax: 314 286 1810
Final: est@wastson.wustl.edu
This clone is available royalty-free through LLNL; contact the
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 274.
                                                                                                 2060 aagaagaaacagcagaaagcagtctcaggttcaacttgccattcaggtgttacttgtgga 2119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11..407
/Organism="Homo sapiens"
/Organism="Homo sapiens"
/Ob_xref="taxon:9606"
/Clone="IMAGE:592696"
/Clone="IMBGE:592696"
/lab_host="SOLR cells (kanamycin resistant)"
  Length 849;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 407;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3.1%; Score 70; DB 9; Le
100.0%; Pred. No. 4.1e-18;
Live 0; Mismatches 0;
3.6%; Score 81; DB 10; I
100.0%; Pred. No. 9.1e-23;
tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       114 9
                                                                                                                                                                                                     2120 taagcacaaatatttcctgtt 2140
                                                                                                                                                                                                                                61 TAAGCACAAATATTTCCTGTT 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AA159952.1 GI:1734443
                                                  81; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
Matches 70; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              143 teccatecat 152
                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οy
```

Search completed: July 23, 2002, 17:46:44 Job time: 9520 sec

ı		·	
			·
			•

```
096691 homo sapien
099284 homo sapien
099284 homo sapien
09609 homo sapien
096093 rhizobium l
096066 mus musculu
09681 escherichia
016138 heliocidari
025384 hemicentrot
01887 caenorhabdi
092285 rhizobium m
091580 prevotella
                                                                                                                                                                                     093998 microscilla
096ke0 pasteurella
018924 caenorhabdi
098495 chlamydomon
094198 homo sapien
065931 mycobacteri
099ve24 drosophila
091799 pseudomonas
021796 microscilla
091799 pseudomonas
02176 mycobacteri
095778 mycobacteri
006776 mycobacteri
0911e5 pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                             093p97 microscilla
090xb6 coturnix cc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.

MEDLINE-20233298; PubMed=10772344;

Wittstock U., Fischer M., Svendsen I., Halkier B.A.;

Roman snail, Helix pomatia.";

HUBMB Life 49:71-76(2000).

EMBL, ARF09924; AAF30402.1; -.

HSSP; PIS848; IFSU.

InterPro; IFR000917; Sulfatase.

Pfam; PF00884; Sulfatase, 1.

PROSITE; PS00149; Sulfatase, 1.

PROSITE; PS00149; Sulfatase, 2.

SEQUENCE 503 AA; 55821 MW; AC45334BA694413D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Helix pomatia (Roman snail) (Edible snail).
Bukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
Helicacea; Helicidae; Helix.
NCBL_TaxID-6536;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      37 CLLTYGYLSWGQALEEEEEGALLAQAGEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGS 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      51;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 503;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           tch 42.9%; Score 1293.5; DB 5; Length al Similarity 48.3%; Pred. No. 8.7e-102; 256; Conservative 82; Mismatches 141; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                        ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               503
                                                                                                                                                          Q9L5W0
P95059
Q93P98
Q9CKE0
Q18924
Q9ATG5
Q9UHJ8
Q65931
                                                                                                                                                                                                                                                                                         Q93P96
Q91799
Q21376
Q95R73
                                                                                                                                                                                                                                                                                                                                                                006776
0911E5
093P97
Q9L4Y2
Q96EG1
Q9Y2K4
Q96CJ0
Q98BQ3
                                                                                                                                            Q922H5
                                                                                    09F821
016138
025384
018837
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                            16
                                                                                                                                                                         16
                                                                                                                                                                                                    16
                                                                                                                                                                                                                                 10
                                                                                                                                                                                                                                                   16
                                                                                                                                                                                                                                                                                                        16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-OCT-2000 (TrEMBLrel. 15, 01-OCT-2000 (TrEMBLrel. 15, 01-DEC-2001 (TrEMBLrel. 19,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
 2525577
2525577
252577
252577
25257
25257
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
2527
 SULFATASE 1.
SULF1.
                                                                                                                                                                                                                                                              243.5
236
230
221.5
219.5
219.5
218
213.5
207
206.5
 391
384.5
383
378
377
371
360
355.5
344.5
344.5
344.5
314.5
314.5
319.5
269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              09NJU8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             90CN90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
Q9NJU8
   අ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9vvml drosophila
Q9h5d8 homo sapien
Q9a22 caulobacter
Q91hk9 mus musculu
Q99ku8 mus musculu
Q96vv7 schizosacch
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9c0v7 schizosacch
Q92uc0 rhizobium m
Q92yf6 rhizobium m
Q9x759 klebsiella
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q9vvm4 drosophila
Q9vvk9 drosophila
Q9nJu7 hellix pomat
Q9vvm1 drosophila
Q9h5d8 homo saplen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9d3b4 mus musculu
Q914y4 klebsiella
                                                                                                                 (without alignments)
714.318 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9nju8 helix pomat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9v6f2 drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Description
                                                                                                                                                                                    MAPRGCAGHPPPPSPQACVC......VPVRYPPKDPRSNPRLNGGV
                                                                                                July 23, 2002, 17:46:49; Search time 133.2 Seconds
               GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                            562222 segs, 172994929 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUMMARIES
                                                                                                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                       OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09VVM1
09H5D8
09A922
09JHK9
099KU8
09C0V7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q92YF6
Q9X759
Q9D3B4
Q9L4Y4
                                                                                                                                                                                                                  BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9V6F2
Q9VVM4
Q9VVK9
Q9NJU7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    90708
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sp_human:*
sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sp_plant:*
sp_rodent:*
sp_virus:*
sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_mhc:*
sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                                                                                                                      sp_archea:*
sp_bacteria:*
sp_fungi:*
                                                                                                                                                                                                                                                                                                                      length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16
                                                                                                                                                 US-09-495-823-7
3012
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB
                                                                                                                                                                                                                                                                                                                                                                                                                         SPTREMBL_19:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                579
996
2266
2266
2266
521
521
537
537
537
537
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    42.9
330.5
330.5
228.1
211.5
211.5
118.7
118.3
113.1
113.0
113.0
113.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 55...
1111...
112...
114...
115...
                                                                                                                                                                                                                                                                                                                   Minimum DB seq
Maximum DB seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1293.5
846.52
846.52
646.53
65.53
65.53
65.53
1303
393
393
393
393
393
393
393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score
                                                                                                                                                                       Perfect score:
Sequence:
                                                                                                                                                                                                                   Scoring table:
                                                                                                                                                                                                                                                            Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                           Database
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Result
No.
                                                                                                  Run on:
```

~

```
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Adams M.D., Celniker S.E., Holt R.A., Faskins R.A., Galle R.F.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfelifer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Rabeson K.Y., Bencas P.V., Berman B.P., Bhandari D., Bolshakov S.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Deasley E.M.,
Ballew R.M., Cabby B.D., Berman B.P., Brocktein P., Brottier P.,
Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
Borkova D., Botchan M.R., Buller H., Cadieu E., Center A., Chandra I.,
A Borkova D., Botchan W.R., Bulke C., Davengort L.B., Davies P.,
A de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
A de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
A dodon K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C.,
Braris N.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
R. Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
R. Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
R. Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Liu X., Matteil B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
A Morkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy L., Murphy L., Musny D.M., Nelson D.L.,
                                                                                                                                                                                                                                                                                                                                                                                                                                         Drosophila melanogaster (Fruit fly).
Eukaryota, Metazoa, Arthropoda, Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                 HSPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDN 328
                                                                                                         SPILADKLKESGYATHWVGKWHLGFYKQEYLPWNRGFDTYFGYLNAAEDYFNHNVPWRQV
                                                                                                                                                                                                                                                   GGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLISLA
                                                                                                                                                                                                                                                                                                                                  EGQIDEDIQLDGYDIWETISEGLRSPRVDILHNID------PIYTKAKNGSWAAGYGI
                                                                                                                                                                                                                                                                                                                                                        WNTAIQSAIRVQHWKLLTGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGKSVWLFNITA
  NATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDTFFGSLLGSGDYYTH-----
                                                                               -YKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQILASHNPTKPIFLYIAYQAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 501 DPYERVDLSNRYPGIVKKLLRRLSQFNKTAVPVRYPPKDPRSNPRLNGGV 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-NAY-2000 (TrEMBLrel. 13, Created)
01-NAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
CG8646 PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ephydroidea; Dro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-BERKELEY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9V6F2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ~
                                                                                                                                                                                                                                                                                                                                                                                                                                                        394
157
                                      115
                                                                               210
                                                                                                                      175
                                                                                                                                                                 569
                                                                                                                                                                                                        225
                                                                                                                                                                                                                                                                                       285
                                                                                                                                                                                                                                                                                                                                389
                                                                                                                                                                                                                                                                                                                                                                         345
                                                                                                                                                                                                                                                                                                                                                                                                                  441
                                                                                                                                                                                                                                                   329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
09V6F2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                    a
                                                                               ð
                                                                                                                      g
                                                                                                                                                                 ò
                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
```

```
Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reanington K., Sanders R.D.C., Scheeler F., Shen H., Shue B.C., Siden Kiamos I., Simpson M., Skroski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Sarman D.A., Weinstock G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Mu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zhong K.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., The genome sequence of Drosophila melanogaster.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDTFFGSLLGSGDYYTHYKC 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ----PLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSS 326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PTLLEAAGGQPALSNLSKQIDGQSIWRALVQDKASPRLNVLHNIDDIWGSAALSVGDWKL 380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   I-----TSSTGKSV-----WLFNITADPYERVDLSNRYPGIVKKLLRRLSQF 526
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            94 HGS-EIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITGKYQIHTGLQHSIIRPTQPNC 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNGGQPTAG-----GSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWY 381
                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                              34 WILCLLTYGYLSWGQALEEEEEGALLAQAGEKLEPSTTSTSQPHLIFILADDQGFRDVGY 93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   213 DSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQILASHNPTK-PIFLYIAYQAVHS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----NGSWAAGYGIWNTAIQSAIRVQHWKLLTGNPGYSDWVPPQSFSNLGPNRWHNER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PTLISLAEGO---IDEDIQLDGYDIWETISEGLRSPRVDILHNIDPIYTKAK-----
                                                                                                                                                                                                                                                                                                                                                                                                 98;
                                                                                                                                                                                                                                                                                                                                                                  Length 542;
                                                                                                                                                                                                                                                                                                                                                               30.5%; Score 920; DB 5; Length 54 38.1%; Pred. No. 7.3e-70; Live 79; Mismatches 169; Indels
                                                                                                                                                                                                                                   E19Fase; FBqn0033763; CG8646.
InterPro; IPR000917; Sulfatase.
Pfam; PF00884; Sulfatase; 1.
PROSITE; PS000523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
SEQUENCE 542 AA; 60278 MW; EACD9C14DD2B3204 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          579 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9VVM4;
01-MAY-2000 (TrEMBLrel. 13, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                        EMBL; AE003821; AAF58475.1; -. HSSP; P15848; 1FSU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NKTAVPVRYPPKDPRSNPR 545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
Matches 213; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q9VVM4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 A 1
      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ОЪ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ō
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

15;

```
10.1-WF-2000 (TERBELFEL. 13, Last sequence update)

DF 01.30FROTEIN.

GRADOR RECTEIN.

GRADOR RECTEIN.

GRADOR RELEASE.

GRADOR RELE
```

14; 192 GFDTFFGSLLGSGDYYTHY--KCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQI 249 73 TSQPHLIFILADDQGFRDVGYHGS-EIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFIT 131 GKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRR 191 Gaps 25 STKPNIVIILIDDMGMNDVSFHGSNQILTPNIDALAYNGILLNKHYVPNLCTPSRATLLT 57; DB 5; Length 579; Indels Pred. No. 1.5e-63; ; Mismatches 192; 28.1%; Score 846.5; 36.5%; Pred. No. 1.5 85; Conservative Query Match Best Local Similarity Matches 192; Conserv 132 85 ò 셤 ö 셤 ö

RC STRAINBERKELER;

KR. SELGENCE, FROW N.A.

RADLINE=20196006; PubMed=10731132;

RADLINE R.A. Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RADLINE R.A., Rogers Y. F., Yandell M.D., Zhang O., Chen L.X.,

RADLINE R.A., Bards A., Bardser B.G., Helt G., Nelson C.R., Miklos G.L.G.,

RADLINE R.A., Banos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RADLIS R.C., Busam D.A., Berler H., Cadieu E., Center A., Chandra I.,

RADLIS R.A., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RADLIS R.D., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RADLIS R., Camley S., Dahlke C., Davenport L.B., Davies P.,

RADLIS R., Camley S., Dahlke C., Davenport L.B., Davies P.,

RADLIS R., Camley S., Dahlke C., Perraz C., Ferriera S., Fleischmann W.,

RADLIS R., Camley B., Gorrell J.H., Gu Z., Guan P., Harris M.,

RADLIS R., Rodira C., Garrell J.H., Gu Z., Guan P., Harris M.,

RADLIS R., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RADLIS R., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum R.A.

RADLIS R., Kodira C.D., Kraft C., Molles R., Molles D., Lin X.,

RADLIS R., Mattei B., McIntosh T.C., McLeod M.P., Mosherson D.,

RADLIS M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RADLIS R., Rodira C.D., Rad., Nison R., Nusskern D.R., Pacleb J.M.,

RADLIS R., Relson K.A., Nison K., Nusskern D.R., Pacleb J.M.,

RADLIS R., Relson K.A., Nison K., Nusskern D.R., Pacleb J.M.,

RADLIS R., Relson K.A., Nison R., Nusskern D.R., Pacleb J.M.,

RADLIS R., Relson K.A., Nison R., Nusskern D.R., Pacleb J.M.,

Relson D.R., Nelson K.A., Nison R., Pollard J., Pull V., Reese M.G.,

Relson R., Relson R.A., Nison R., Pollard J., Pull V., Reese M.G.,

Relson R., Relson R.A., Nison R., Pull V., Pull V., Reese M.G.,

Relson R., Relson R.A 322 381 442 LLGNRGLTKDRIRQMRSEATETCPPIEGONPLESHFKCEPL----KAPCFPDLAKDPCE 496 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pteryyota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila; NCBL\_TaxID=7227; 145 GFDHHFGYYNGYIDYYDHQVRMLDRNYSAGLD-FRRDLEPCPEANGTYATEAFTSEAKRI 203 419 LHNIDPIY-----TKAKNGSWAAG-YGIW------NTAIQSAIRVQHWKL 456 323 WSPLLKERGYVSNQAIHAVDWLPTLAGAAGVSLPQDLPLDGINLWPMLS-GNEEPKPRTM 305 NNVTLALKTYGFYNNSIIIYSSDNGGQPTAG----GSNWPLRGSKGTYWEGGIRAVGFV 360 HSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRV-DI LTGNPGY-----SDWVPPQSFSNLGPNRWHNERITSSTGKSVWLFNITADPYE 382 IHVLDEVFGYSSYMRDTLKYVNGSSFKGRYDQWLGELETNEDDPLGESYEQHVLASDVQS 250 LASHNPTKPIFLYIAYQAVH----SPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAI 505 RVDLSNRYPGIVKKLLRRLSQFNKTAVP-VRYPPKDPRSNPRLNGG 549 Last sequence update) Last annotation update) 966 Created) Drosophila melanogaster (Fruit fly) PRT; 01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-JUN-2001 (TrEMBLrel. 17, CG5584 PROTEIN. PRELIMINARY; SEQUENCE FROM N.A. CG5584. 09VVK9; Q9VVK9 457 497 RESULT **09VVK9** g g ò g ò a ò g ò δ g Q

```
Adams M.D., Celniker S.E., Holf R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfelffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikkos G.L.G., Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.
                           Helix pomatia (Roman snail) (Edible snail).
Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
Helicacea; Helicidae; Helix.
NCBI_TaxID=6536;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 195 TFFGSLLGSGDYYTHYKCD-----SP----SP-----GMCGYDLYEN 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SYYGYLEGGEDYYTYYNCDTFHNRTTPADTSILESYSPKNILLGKHEDENKWGGYDLKDM 206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa, Arthropoda; Tracheata; Hexapoda; Insecta;
Pletygota; Neoptera, Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidea; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              75 QPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITGKY 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              135 QIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFD 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-20233298; PubMed=10772344;
MEDLINE-20233298; PubMed=10772344;
Wittstock U., Fischer M., Svendsen I., Halkier B.A.;
"Cloning and characterization of two cDNAs encoding sulfatases in t ROMAn snail, Helix pometia.";
IUBMB Life 49:71-76(2000).
EMBL, AF109925; AAF30403.1; -.
HSSP, PISH848; IFSU.
InterPro; IPR000917; Sulfatase.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS0049; SULFATASE_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 226 DNAAWDYDNGIYSTQMYTQRVQQIL-ASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     56; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2F877BB1737A9C49 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21.5%; Score 646.5; DB 5
51.0%; Pred. No. 5.4e-47;
ive 31; Mismatches 56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-BERKELEY;
MEDLINE-20196006; Pubmed-10731132;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30391 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-JUN-2001 (TrEMBLrel. 17,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 51.09
Matches 123; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          266 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ephydroidea; Dros
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CG7408 PROTEIN.
CG7408.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  285 I 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            266 i 266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9VVM1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9VVM1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   009 VVW

100 VVW

100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ω
                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             18;
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zhong X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., The genome sequence of Drosophila melanogaster.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           88 FRDVGYHGS-EIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITGKYQIHTGLQHSIIR 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 : :! ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | ::: | :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PIFLYIAYQAVHS-----PLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLALK 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRV-.DILHNIDPI 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ------DVWRL---SALQMGQWKYVNGTTASGRYDSVLTYRELDDLDPRDSRY 432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               147 PTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDTFFGSLLGSGDY 206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               207 Y-----THYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQILASH-NPTK 257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               313 TYGFYNNSIIIYSSDNGGQPTAG----GSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNK 367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----SVWLFNITADPY 503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 433 AVTVRNSATSRALSRYDLRRLTQQRISLTRRLAAVRCGDLQRSCNPLLEECLYDILSDPC 492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28 GALAGFWILCLITYGYLSWGQALEEEEEGALLAQAGEKLEPSTTSTSQPHLIFILADDQG 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | | :|| | :|| | 3 GLSAFLLLCL------QRPVKSDESAA------ARRPNIIIMADDMG 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | || : | | || : | || || : | : || || || : || 39 FDDVSFRGRELLPRIDALAYHGRLLDRLYAPAMCTPSRGALLSGRYPIHTGTQHFVIS 98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               426 YTKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGNPG---YSDWVPPQSFSNLGP----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; Score 752; DB 5; Length 996;
; Pred. No. 4e-55;
98; Mismatches 188; Indels 116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C92AA3788D082FB3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         504 ERVDL--SNRYPGIVKKLLRRLSQFNKTAVPVRYPPKDPRSNPRLN 547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                             F19Base; FBG00036755; CG5584.
InterPro; IPR000917; Sulfatase.
Pfam; PF00804; Sulfatase. 1.
PROSTIE; PS00149; SULFATASE.; 1.
SEQUENCE 996 AA; II13514 MW; C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (TrEMBLrel. 15, C
(TrEMBLrel. 15, I
(TrEMBLrel. 19, I
                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AE003522; AAF49302.1;
HSSP; P15848; 1FSU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 25.0%
Best Local Similarity 31.4%
Matches 184; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-OCT-2000 (
01-DEC-2001 (
SULFATASE 2 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-OCT-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q9NJU7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q9NJU7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
Q9NJU7
                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       84448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

4

31;

Length 266;

S

```
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Cherry J.W., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA de Pablos B., Delcher A., Deng S., Manar Rocha S., Dunkov B.C., Dunn P.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Glodek A., Gong F., Gorg C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Melson D.R., Nelson K., Nixon K., Nixon D.R., Pecleb J.M.,
RA Ralazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Spier E., Spradling A.C., Stapleton M., Stupski M.P., Wang X.,
RA Shue B.C., Siden Kiamed I., Singson M., Stupski M.P., Wang S., Yao Q.A.,
RA Shen S.M., Woodage T., Weinschock G.M., Weissenbach J.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhang G., Zhao Q., Zheng L.,
Rh Genome sequence of Drosophila melanogaster.",
Rh Genome sequence of Drosophila melanogaster.",
Rh Grans Remai, 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               200 LLGSGDYYTH-YKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQILASH---NP 255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                256 TKPIFLYIAYQAVHS-----PLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLA 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LKTYGFYNNSIIIYSSDNGGQPTAG----GSNWPLRGSKGTYWEGGIRAVGFVHSPLLK 365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----IYTKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTG--NPG-YSDWVPPQSFSNL 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       299 VAEPHLSYTRGK -------335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       140 LQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDTFFGS 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           366 NKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDP- 424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   336 DPRAVEYEELVRNT -- SVWLOLQOVSFGERNISELRDQSRIECPDPATGVKPCLPLEGPC 393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SQPLFLLLNHLAPHAANDDDPMQAPAEEVSRFEYISNKTHRYYAAMVSRLDKSVIDA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LFNITADPYERVDLSNRYPG--IVKKLLRRLSQFNKTAVPVRYPPKDPRSNPR 545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18.5%; Score 556; DB 5; Length 486; 31.3%; Pred. No. 7.1e-39;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       65; Mismatches 166; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF00884; Sulfatase; 1.
SEQUENCE 486 AA; 55063 MW; 4327F6726A98F524 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      476 GPNRWHNERITSSTGKSVW--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR000917; Sulfatase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              :lyBase; FBgn0036765; CG7408.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 148;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  394
              ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
```

```
ö
                                                                                                                                                                                                                                      Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T., Yamada K., Fujiu Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.; Subhitted (Aug. 2000) to the EMBL/GenBank/DDBJ databases.

Submitted (Aug. 2000) to the EMBL/GenBank/DDBJ databases.

HSSP; P15848; 1FSU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN—ATCC 19089 / CB15;

STRAIN—ATCC 19089 / CB15;

MIDLINE—2117698; bubmed=11259647;

Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E., Bisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R., Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B., DeBoy Y.T., Dodson R.J., Durklin A.S., Gwinn M.L., Haft D.H., Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K., Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O., Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.; "Complete genome sequence of Caulobacter crescentus."; Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       455 KLLTGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGKSVWLFNITADPYERVDLSNRYPG 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Bacteria; Proteobacteria; alpha subdivision; Caulobacter group; Caulobacter.
                                                                                                                                                                                                                                                                                                                                                                                                                                       16.8%; Score 507; DB 4; Length 169; 96.9%; Pred. No. 2.2e-35; Live 1; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                         169 AA; 18905 MW; 4583C92FFFCC3380 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
Last sequence update)
Last annotation update)
                                                                             Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    515 IVKKLLRRLSQFNKTAVPVRYPPKDPRSNPRLNGGV 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     521 AA.
                                                                Created)
                                                                                                              CDNA: FLJ23548 FIS, CLONE LNG08487.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Interpro; IPR000917; Sulfatase.
Pfam; PF00884; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (TrEMBLrel. 19,
                                                           01-MAR-2001 (TrEMBLrel. 16, 01-MAR-2001 (TrEMBLrel. 16, 01-DEC-2001 (TrEMBLrel. 19)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (TrEMBLrel. 17,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       93; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SULFATASE FAMILY PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Caulobacter crescentus
                                                                                                                             (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HSSP; P15848; 1FSU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=69394;
                                                                                                                                                                            NCBI_TaxID=9606;
                                                                                                                             Homo sapiens
                                                                                                                                                                                                                             TISSUE=LUNG;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9A922;
01-JUN-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-JUN-2001
01-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TIGR; CC1172
                                                                                                                                                                                                                                                                                                                                                                                         SEOUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local
                                             Q9H5D8;
                              Q9H5D8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09A922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
RESULT
Q9H5D8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09A922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
```

```
Q99KU8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q99KU8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          10
                                                                                                                                                                                                                                                                                                                                                                           88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
 g
                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Qγ
                                                                                                                                                                                                                                                                                                                                                                          g
                                                                               16;
                                                                                                                                                                                                                             249 EAVKAIDA-NKNRPFFEMYLAYNAVHTPLQAPKADYDALSHIKDHRMRVYAAMVRNLDRNV 307
                                                                                                                                                                                                                                                                                                                                                                                                    367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DP--HKAI-------PWRSGGYKTVLAGGWKLQVAKE--------PNK--- 449
                                                                                                                                                                                                                                                                                             -----PGVV-----ESRQDWDPIDKFLWGAAPFAVQFNGGKLFNPSHYMTDYLTD 248
                                                                                                                                                                                                                                                                                                                                                                           NNVTLALKTYGFYNNSIIIYSSDNGGQPTAG--GSNWPLRGSKGTYWEGGIRAVGFVHSP 362
                                                                                                                                                                                                                                                                                                                                                                                                                                                         368 EQLPAGAVYRSPVGHVDIFATAAGAAGAPTPKDRVLDGVDL-----VPFVKGQSSG 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DPIYTKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGNPGYSDWVPPQSFSNLGPNRWHN 482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PETRAIN-129/SVJ, AND C57BLACK;
MEDLINE-20164110; PubMed-10699374;
MEDLINE-20164110; PubMed-10699374;
MOTAGO A., Yamagishi A., Tomatsu S., Fukuda S., Copeland N.G.,
MOTAGO A., Yamagishi A., Tomatsu S., Fukuda S., Copeland N.G.,
Orii K.E., Isogai K., Yamada N., Kato Z.I., Jenkins N.A.,
Gilbert D.J., Sukegawa K., Orii T., Kondo N.;
"The mouse N-acetylgalactosamine-6-sulfate sulfatase (Galns) gene:
cDNA isolation, genomic characterization, chromosomal assignment and
analysis of the 5'-flanking region.";
Biochim. Biophys. Acta 1500:323-334(2000).
EMBL; AF112230; AAF63858.1;
EMBL; AF112231; AAF63858.1;
JOINED.
EMBL; AF112231; AAF63858.1;
JOINED.
EMBL; AF112232; AAF63858.1;
JOINED.
                                                                                                        68 PSTTSTSQPHLIFILADDQGFRDV----GYHGSEIKTPTLDKLAAEGVKLEN-YYVQPI 121
                                                                                                                                                            CTPSRSQFITGKYQIHTGLQHS------151
                                                                                                                                                                                    CAPSRAAIMTGRYATREGFEFTPTPVAFSRVVGGHAGDPLHPSRFNQAEVKNMPKDENVL 144
                                                                               Gaps
                                                                                                                        CLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDTFFGSLLGS-----GD
                                                                                                                                                                                                                                                                    206 YYTHYKCDSPGMCGYDLYENDNAAWD------YDNGIYSTQMYTQ
                                                                                                                                                                                                                                                                                                                                                                                          Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                        245 RVQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAI
                                                                                                                                                                                                                                                                                                                                                                                                                                363 LLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      483 ERITSSTGKSVWLFNITADPYERVDLSNRYPGIVKKLLRRLSQFNKTAVPVRYP 536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    450 ------TWLFDLSTDPTERQELSKARPEKLREMQAIMAQLDGQMMKPTWP 493
                                                                             Indels 130;
                                                    Length 521;
             DCD14EEFD090E768 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
N-ACFTYLGALACTOSAMINE-6-SULFATE SULFATASE.
                                                   15.7%; Score 472; DB 16;
26.6%; Pred. No. 1.2e-31;
tive 69; Mismatches 193;
             MΨ,
             56709
                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
             521 AA;
                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
  proteome
                                                                  Best Local Sim
Matches 142;
Complete p
                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             09JHK9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              09ЛНК9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   423
                                                                                                                                 25
                                                                                                                                                                                                                152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            419
                                                                                                                                                           122
                                                                                                                                                                                     85
                                                                                                                                                                                                                                                                                             204
                                                                                                                                                                                                                                                                                                                                                                           305
                                                                                                                                                                                                                                                                                                                                                                                                     308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
 SE
                                                                                                                                 g
                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                            ò
                                                                                                                                                                                    g
                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LPIRNGFYTTNAHARNAYTPQEIMGGIPNSEHLLPELLKKAGYTNKIVGKWHLG-HRPQF 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   134 YQIHTGL-----QHSIIRPTQ-PNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKEC 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     147 HPLKHGFNEWFG------SPN-CHFGPYDNKAKPNIPVYRDWEMVGRFYEE 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----TQMYTQR-VQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRSI 285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                341 LRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDG 400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                401 YDIWETISEGLRSPRVDILHNIDPIYTKAKN-----GSWAAGYGIWNTAIQSAIRVQH 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WEEFTQG---TDFCPGQNVSG-----VTTHTQEEHTELPLIFHLGRDPGERFPLS 453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   76 PHLIFILADDQGFRDVGYHGSEIK-TPTLDKLAAEGVKLENYY-VQPICTPSRSQFITGK 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28 PNIVLLIMDDMGWGDLGVNGEPSRETPNLDRMAAEGMLFPSFYSANPLCSPSRAALLTGR 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 191 FPINRKTGEANLTQLYTQEALDFIQTQHARQSPFFLYWAIDATHAPVYAS-----RQF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             304 FLCGKQTTFEGGMREPAIAWWPGHIAAGQVSHQLGSIMDLFTTSLSLAGLKPPSDRVIDG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Craniata; Vertebrata; Euteleostomi;
Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   187 MPTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNAA------WDYDNGIYS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       286 ININ-RRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNG----GQPTAGGSNWP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LDLLPTMLKGQMMDR-----PIFYYRGNTLMAVTLGQYKAHLWTWTNS------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WKLLTGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGKS----VWLFNITADPYERVDL-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels 112;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 520;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-JUN-2001 (TrEMBLrel. 19, Last annotation update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
SIMILAR TO GALACTOSAMINE (N-ACETVL)-6-SULFATE SULFATASE.
                                                                                                                                                                                                                                                                                                                                                             MOST, MISSONS Galns.
InterPro; IPRO00917; Sulfatase.
Pfam; PF00884; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00523; SULFATASE_1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       --SNRYPGIVKKLLRRLSQFNKTAVPVRYPPKDPRSNPRLN 547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     454 FHSDEYQDALSRTTQVVQEHQKSLVP-----GQPQLN 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 81; Mismatches 190;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               14.8%; Score 447; DB 11; 26.5%; Pred. No. 1.6e-29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           520 AA
EMBL, AF112233, AAF63858.1; JOINED.
EMBL, AF112234; AAF63888.1; JOINED.
EMBL, AF112235, AAF63858.1; JOINED.
EMBL, AF112236; AAF63858.1; JOINED.
EMBL, AF112237; AAF63858.1; JOINED.
EMBL, AF112239; AAF63858.1; JOINED.
EMBL, AF112239; AAF63858.1; JOINED.
EMBL, AF112240; AAF63858.1; JOINED.
EMBL, AF112244; AAF63858.1; JOINED.
EMBL, AF112244; AAF63858.1; JOINED.
EMBL, AF11244; AAF63858.1; JOINED.
EMBL, AF11244; AAF63858.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Metazoa; Chordata;
Mammalla; Eutherla; Rodentla;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best_Local Similarity 26.59
Matches 138; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mus musculus (Mouse).
                                                                                                                                                                                                                                                                                                         EMBL; AF111346; AAF
HSSP; P15848; 1FSU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
```

19;

```
STRAIN=1021;
MEDLINE=21395608; PubMed=11481431;
Finan T.M. Weidner S., Wong K., Buhrmester J., Chain P.,
Vorhoelter F.J. Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
Golding B., Puehler A.,
"The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
fixing endosymbiont Sinorhizobium meliloti.";
Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
                                                                                                                                                                                                                                                                                                          TDNHIAGLGQMAETVRRFSKVWGGKPGYEGYLNDRVAALPEILQEAGYYTTMSGKWHLGL 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LAE----GQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPIYTKAKNGSWAAGYGIWN 442
                                                                                                                                                                                                                                                                                                                                                                                                          -TPDRYPSKRGFKESFALLPGGGNHFAY ---- EPGTRENPAVPFLPPLYTHNHDPVDHKS 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         286 ININRRR-----YAMLSCLDEA 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LRKNRLQAQKDLGLIPENVIPAPVDGMGTKSWDELTTEEKEFSARTMEVYAAMVELLDLN 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 --GGSNW-----PLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLIS 386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         364 IWYGPRWAQAATAPSRLSKGFITEGGIRCPAIIRYPPLIKPDIISDEFVTVMDILPTILE 423
                                                                                                                                                                                               73 TSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITG 132
                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                                    YRKECMPTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCGY-----DLYENDNAAWDYDN 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TAIQSAIRVQHWKLLTGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGKSVW-LFNITAD 501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              471 --- ORAIRKGNYKAI------YVPKEGI-------KTEWELYDLSQD 501
                                                                                                                                                     Gaps
                                                                                                                                                                                                                                       9 SKKPNFLVIVADDLGWSDVSPFGSEIHTPNIERLAKEGVRLTNFHTASACSPTRSMLLSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LKNFYSSNYFAEKLIDQLKNREKSQSFFAYLPFTAPHWPLQSPKEYINKYRGRYSEGPDV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       133 KYQIHTGL-----QHSII---RPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGF
                                                                                                                                                                                                                                                                                                                                                                                                                                                       --GIYSTQMYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYR-----SI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
PUTATIVE ARYLSULFATASE PROTEIN (EC 3.1.6.1).
Rhizobium meliloti (Sinorhizobium meliloti).
Plasmid psymb (megaplasmid 2).
Bacteria: Proteobacteria: alpha subdivision; Rhizobiaceae group;
                                                                                                        ; Score 432; DB 3; Length 554;
; Pred. No. 3.3e-28;
82; Mismatches 178; Indels 158;
Pfam; PF00884; Sulfatase; 1.
PROSITE; PS00149; SULFATASE_2; 1.
SEQUENCE 554 AA; 62422 MW; 2BBF422DEDOC5BA7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           537 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INNVTLALKTYGFYNNSIIIYSSDNGGQPTA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                502 KGELENLAKVHPDILNELI 520
                                                                                                           14.38;
                                                                                                      Query Match
Best Local Similarity 25.2*
Matches 141; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          502 PYERVDLSNRYPGIVKKLL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=382;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q92UC0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    244
                                                                                                                                                                                                                                                                                                                          69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q92UC0
    DR
DR
SQ
                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                            21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   400
                                                                                                                                                                                                                                                                                                                                                                                                               147 HPLKHGFDEWFG-----SPN-CHFGPYDNKAKPNIPVYRDWEMVGRFYEE 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              286 ININ-RRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNG----GQPTAGGSNWP 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      453
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                76 PHLIFILADDQGFRDVGYHGSEIK-TPTLDKLAAEGVKLENYY-VQPICTPSRSQFITGK 133
                                                                                                                                                                                                                                                                                                                                                                                    134 YQIHTGL-----QHSIIRPTQ-PNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKEC 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MPTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNAA------WDYDNGIYS- 238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   191 FPINRKTGEANLTQLYLQEALDFIRTQHARQGPFFLYWAIDATHAPVYAS-----RQF 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FLCGKQTTFEGGMREPAIAWWPGHIAAGQVSHQLGSIMDLFTTSLSLAGLKPPSDRVIDG 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WKLLTGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGKS----VWLFNITADPYERVDL- 508
                                                                                                                                                                                                                                                            Indels 112; Gaps
                                                                                                                                                                                                                                                                                                                            ------TQMYTQR-VQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRSI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     401 YDIWETISEGLRSPRVDILHNIDPIYTRAKN-----GSWAAGYGIWNTAIQSAIRVQH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LDLLPTMLKGQMMDR-----PIFYYRGNTLMAVTLGQYKAHLWTWTNS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   341 LRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      O'Neil S., Harris D., Wood V., Rajandream M.A., Barrell B.G.; Submitted (Apr. 2001) to the EMBL/GenBank/DDBJ databases. EMBL; AL590604; CAC36911.1; -. HSSP; PIS848; 1FSU.
                                                                                                                                                                                                                 Length 520;
                      the EMBL/GenBank/DDBJ databases
                                                                              InterPro; IPR000917; Sulfatase.
Pfam. PF00884; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
SEQUENCE 520 AA; 57738 MW; 3ADE9662B800F745 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Schizosaccharomyces pombe (Fission yeast).

Bukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
Schizosaccharomyces.

NCBI_TaxID=4896;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            454 FHSDEYQDALSRTTQVVQEHQKSLVP-----GOPOLN 485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --SNRYPGIVKKLLRRLSQFNKTAVPVRYPPKDPRSNPRLN 547
                                                                                                                                                                                                             14.7%; Score 443; DB 11; 26.5%; Pred. No. 3.5e-29; ive 80; Mismatches 191;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                554 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR000917; Sulfatase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-JUN-2001 (TrEMBLrel. 17, 01-DEC-2001 (TrEMBLrel. 19, PUTATIVE ARYLSULFATASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-JUN-2001 (TrEMBLrel. 17,
                 Submitted (FEB-2001) to the EMBL; BC004002; AAH04002.1; HSSP; P15848; 1FSU.
                                                                                                                                                                                                             Query Match
Best Local Similarity 26.59
Matches 138; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SPBPB10D8.02C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-972H-;
  Strausberg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09C0V7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9C0V7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
                                                                                                                                                                                                                                                                                                                                            28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            509
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  407
```

g

ò

ò a ð

a

ó

셤 ð a

õ

q

셤 ò g

ò

셤

à

363

```
141;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Dierks T.;
     Hydrolase;
SEQUENCE
                                                          Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09X759;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09X759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 14
                                                                                                                                                                                                                                                                                     214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  369
                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             09X759
     SS
                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                       ò
                                                                                                                                                                                                  g
                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                   22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-1021;
MEDLINE-21396509; PubMed-11481432;
MEDLINE-21396509; PubMed-11481432;
MEDLINE-21396509; PubMed-11481432;
MEDLINE-21396509; PubMed-11481432;
Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
Gurjal M., Hong A., Hulzar L., Hyman R.W., Kahn D., Kahn M.L.,
Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
Yeh K.-C., Davis R.W., Federapiel N.A., Long S.R.;
"Nucleotide sequence and predicted functions of the entire
Sloorhizoblum mellloti psymm megaplasmid.";
Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
                                                                                                                                                                                                                                                                                                                                                                        384 DIVAKCLIGYAAGAKTFKVHLDGYNLMPFLSGSSNDAPRRDFLY--------427
                                                                                                                              65 KLEPSTTSTS--QPHLIFILADDQG-FRDVGYHGSEI--KTPTLDKLAAEGVKLENYYVQ 119
                                                                                                                                              PICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHL 179
                                                                                                                                                                                                     OSCTAGRAAFITGQSPIRTGLT-KVGLPGADIGIQPEDATVAELLKSLGYATGQFGKNHL 161
                                                                                                                                                                                                                                                                                                   QMYTQRVQQILASH--NPTKPI-----FLYIAYQAVHSPLQAPGRYFEHYRSI- 285
                                                                                                                                                                                                                                                                                                                           RFGPRGVIKATADGKIEDTGPLSVKRMETVDQEFLEAALDFIDRKAKGDAPWFCYFNSTR 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    384 -----LISLAEGQIDEDIQLDGYDIWETISEGLR-SPRVDILHNIDPIYTKAKNGSWAA 436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GYGIWNTAIQ-SAIRVQHWKLL---TGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGKS 492
                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                467 PLLFNLRSDPFERGDTSMEYEXWFFDRSFVVVPSQAVVAKWLESFKEF----PIRQKP 520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VWLFNITADPYERVDLSNRY------PGIVKKLLRRLSQFNKTAVPVRYPP 537
                                                                                                                                                                                                                                           GFYRKECMPTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYST
                                                                                                                                                                                                                                                                G-DKDEFLPTAHGFDEFFGNLY-----HLNAEE------EPENPDYPQDPA-FRK
                                                                                                                                                                                                                                                                                                                                                                                                              334 A--GGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYPT-----
                                                                                                                                                                                                                                                                                                                                                                                                                                    324 TWPDGGNTPFRGEKATNWEGGFRVPMCIRWPGVIKPGTIHNEPFSHYDLIPTFCAAAGEP
                                                                                                                                                                                                                                                                                                                                                      --ININRRRYAAMLSCL------DEAINNVTLALKTYGFYNNSIIIYSSDNGGQPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----WNDDGELVAVRVQNWKVVFKSQDHEGIGVWRQP--FTEL------RA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Plasmid pSymA (megaplasmid 1).
Bacteria: Proteobacteria: alpha subdivision; Rhizobiaceae group;
Rhizobiaceae; Sinorhizobium.
NCBI_TaxID=382;
                                                                      Query Match 13.1%; Score 395.5; DB 16; Length 537; Best Local Similarity 25.8%; Pred. No. 4.1e-25; Matches 139; Conservative 78; Mismatches 195; Indels 127;
                 1 protein; Complete proteome.
6A090FE52F810220 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
PUTATIVE ARYLSULFATASE (EC 3.1.6.8).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Rhizobium meliloti (Sinorhizobium meliloti).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
EMBL; AL603646; CAC49614.1; -.
Hydrolase; Plasmid; Hypothetical
SEQUENCE 537 AA; 59294 MW; 6P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SMA1683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               092YF6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 13
                                                                                                                                                         43
                                                                                                                                                                                    120
                                                                                                                                                                                                                103
                                                                                                                                                                                                                                                                      162
                                                                                                                                                                                                                                                                                                                            204
                                                                                                                                                                                                                                           180
                                                                                                                                                                                                                                                                                                  240
                                                                                                                                                                                                                                                                                                                                                        286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 092YF6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NW
KW
SO
                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                        g
                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                               ò
                                                                                                                                                                                    à
                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
```

```
24;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                - WF--YFIENELIPGAARVGHYKAVFNLRGDNGQATGGLAVDINLGWKGAQSYVAIVPQ 470
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       432 GSWAAGYGIWNTAIQSAIRVQHWKL---LTGNPGYS------DWVPPQSFSNLGPN 478
                                                                                                                                                                                            75 QPHLIFILADDQGFRDVG-YHGSE---IKTPTLDKLAAEGVKLENYYVQPICTPSRSQFI 130
                                                                                                                                                                                                                                                                                           131 TGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTR 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IIYSSDNGG----QPTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHI 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      378 TDWYPTLISLA-----EGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPIYTKAKN 431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RWHNERITSSTGKSVWLFNITADPYERVDL-SNRY------PGIVKKLLRRLSQFN 527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   311 VFYTTDNGAWQDVYPDAGYT--PFRGTKGTVREGGNRVPAIAFWPGKIQPGSRSHDVVGG
                                                                                                                                                                                                                      -----SPGMCGYDLYEN--DNAAWDYDNGIYSTQMYTQRVQQILASHNPTKPIFL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         256 NVNFMKVHQP-NLPAPEFQHK----SLSKSKYADSVVELDTRIGRILDKLRETGMDKNTL
                                                                                                                                                                                                                                                                                                                                                                                              191 RGFD-----GSLLG--SGDYYTHYKCD-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           262 YIAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        "The iron sulfur protein AtsB is required for posttranslational formation of formylglycine in the Klebsiella sulfatase.";
J. Biol. Chem. 274:15375-15381(1999).
EMBL; AJJ31525; CAB40961.1; -.
InterPro; IPR000917; Sulfatase.
                                                                                           Score 393; DB 16; Length 555;
Pred. No. 7.1e-25;
76; Mismatches 184; Indels 154;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=DSM 681;
MEDLINE-99269066; PubMed-10336424;
Szameit C., Miech C., Balleininger M., Schmidt B., von Figura
  Hypothetical protein; Complete proteome. 61221 MW; 76500C482ACF42F1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
3.1.6.1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bacteria; Proteobacteria; gamma subdivision;
Klebsiella.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                              13.0%;
25.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1999 (TrEMBLrel. 12, 01-NOV-1999 (TrEMBLrel. 12, 01-JUN-2001 (TrEMBLrel. 17, ARYLSULFATASE PRECURSOR (EC
                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----VEYPPRKLQS 520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Klebsiella pneumoniae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 528 KTAVPVRYPPKDPRS
Plasmid; 555 AA;
                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=573;
```

```
Search completed: July 23, 2002, 19:06:40 Job time: 4791 sec
                                                                                                                                               20;
                                                                                                                                                                                                                                                STQMYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQAP-----GRYFEHYRSIININ 289
                                                                                                                                                                                                                                                                                                                                                                 -----YAAMLSCLDEAINNVTLA 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    373 VSNAP-----YANYHKT--TSAQGGINTDFMISGPGITRHGKIDASTMAVY 416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DVAPTLYE------FAGIDPNKSLAKK------PVLPMIGVSFKRYLTGEV 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GYSDWVPPQSFSNLG------PNRWHNERITS---STGKSVW-LFNITADPYERVD 507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             75 QPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITGK- 133
                                                                                                                                                                                                                                 -----YQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFY 182
                                                                                                                                                                                                                                                                                         RKECMPTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCG----YDLYENDNAAWDYDNGIY 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LKTYGFYNNSIIIYSSDNGGQPTAG------GSNWPLRGSKGTYWEGGIRAVG 357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DIWETISEGLRSPRVDILHNIDPIYTKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGNP 461
                                                                                                                                                                                         STRAIN=CS7BL/6J; TISSUE-MEDULLA OBLONGATA;
MEDILINE-21085660; Pubmed-11217851;
Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
                                                                                                                                                                                                                                                                                                                                                                                                                           - PGATPKERGFNHAF-AFMGGGT--SHFNDAIP--LGTVEAFHTYYTRDGERVSLPDDFY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
NCBI_TaxID=10090;
                                                                                                                 Score 393; DB 2; Length 577;
Pred. No. 7.5e-25;
.; Mismatches 188; Indels 162;
                                                           POTENTIAL.
7E897EDB2CABD18C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    525 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
Pfam; PF00884; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
                                                                                                                                            81;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      510 VAAEHPDRVKAMSEAYEAFAK 530
                                                          20 PC
64155 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 508 LSNRYPGIVKKLLRRLSQFNK 528
                                                                                                                 13.0%;
23.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-JUN-2001 (TrEMBLrel. 17, 01-JUN-2001 (TrEMBLrel. 17, 01-DEC-2001 (TrEMBLrel. 19,
                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6330406P08RIK PROTEIN.
6330406P08RIK.
                                                                      577 AA;
                                                                                                                                Similarity
                                         Signal; Hydrolase.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                         RRR----
                                                                                                                 Query Match
Best Local Simi
Matches 130;
                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09D3B4
                                                                                                                                                                                                                                                                                       183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          162
                                                                                                                                                                                                                                 134
                                                                                                                                                                                                                                                                                                                                                                             193
                                                                                                                                                                                                                                                                                                                                                                                                         290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  311
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
  DR
DR
FT
SQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                          ò
                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

```
26;
Radto T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
Radota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Re Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Ruehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Ouackenbush J.,
Rotrim L. M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
Roustein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Nyanahaw-Boxooka K., Wang K.H., Weltz C., Whittaker C., Wilming L.,
Wanshaw-Boxis A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   71 TSTSQPHLIFILADDQGFRDVGYHGSEIK-TPTLDKLAAEGVKLENYY-VQPICTPSRSQ 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            412
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       129 FITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             228 AAWDYDNGIYSTQMYTQR-VQQILASHNPTKPIFLYIAYQAVHSPLQ-APGRYFEHYRSI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   209 IVEQPVNLSGLAQKYAERAVEFIEQASTSGRPFLLYVGQAHMHVPLSVTPPLAHPQRQSL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             286 ININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNGGQPTA----GGSNWP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LRG------SKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHIT-----DMYP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----LFGKSQMGHRVLFHPNSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      435 AAG-YGIWNTAIQSAIRVQHWK--LLTGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRRGFDTFFG----SLLGSGDY--YTHYKCD-----SPGMCGY----DLYENDN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 383 TLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPIYTKAKNG-----SW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----GSVGPEQHHVAPL
                                                                                                                                                                                                                                                                                                    Hayashizaki Y.; "Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13.0%; Score 391.5; DB 11; Length 525; 26.7%; Pred. No. 8.7e-25; ive 89; Mismatches 168; Indels 121;
                                                                                                                                                                                                                                                                                                                                                                EMBL; AK018132; BAb31086.1; -.
HSSP; P15289; 1ANK.
MGD; MGI:1921258; 6330406P08Rik.
InterPro; IPR000917; Sulfatase.
Pfam; PF00884; Sulfatase; 1.
PROSTIE: PS00149; SULFATASE.2; 1.
SEQUENCE 525 AA: 57449 MW; DF8889E39866DIEIF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---IFNLEDAADEGMPLQKGSPEYQEVLQQVTRALA 487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SVWLFNITADPYERVDL - - - SNRYPGIVKKLLRRLS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            371 TVIALAGASLPPNRKFDGRDVSEV-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 13.0%
Best Local Similarity 26.7%
Matches 138; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  341 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  455
        Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
```

```
July 23, 2002, 17:07:14 ; Search time 79.69 Seconds (without alignments) 766.603 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1980.DAT:*
2. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1981.DAT:*
3. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1981.DAT:*
4. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1981.DAT:*
5. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1981.DAT:*
5. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1981.DAT:*
6. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1981.DAT:*
7. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1981.DAT:*
8. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1981.DAT:*
9. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1981.DAT:*
11. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1981.DAT:*
12. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1991.DAT:*
13. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1991.DAT:*
14. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1991.DAT:*
15. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1991.DAT:*
16. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1991.DAT:*
17. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1991.DAT:*
18. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1991.DAT:*
19. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1999.DAT:*
21. (SIDSI) gogdata/geneseq/geneseqp-embl/AA1999.DAT:*
22. (SIDSI) gogdata/geneseqg-embl/AA1999.DAT:*
22. (SIDSI) gogdata/geneseqg-embl/AA1999.DAT:*
                                                                                                                                                                                                                                                                                                                                                                         1 MAPRGCAGHPPPPSPQACVC.....VPVRYPPKDPRSNPRLNGGV 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  747574
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 747574 segs, 111073796 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                      OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                     BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A_Geneseq_032802:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                    US-09-495-823-7
3012
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                        Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Searched:
                                                                                                                                                                                    Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description	Human 23553 sulfat	Human PRO708 prote	Human PRO708 (UNO3	Human PRO polypept	Human 25278 sulfat	Human novel protei	Human sulfatase pr	Human novel protei	Human protein segu	Drosophila melanog	Drosophila melanog
Ð	AAB85483	AAY41701	AAB44257	AAU29061	AAB85482	AAU14421	AAB51184	AAU14185	AAM25415	ABB63962	ABB63246
DB	22	20	21	22	22	22	22	22	22	22	22
Query Match Length DB	599	515	515	515	569	289	533	366	187	542	579
Query						51.9					
Score	2996	2307.5	2307.5	2307.5	1859.5	1562	1415.5	1079	686	920	846.5
Result No.	-	7	m	4	S	9	7	80	6	10	11

Drosophila melanog Novel human enzyme Drosophila melanog Escherichia coli p Fam consensus seq Human protein SEQ Human polypeptide Human steroid sulp Human sulfatase pr Human sulfatase pr Human polypeptide Novel human diagno Human polypeptide Human polypeptide Novel human diagno Sea urchin sulfatase pr Human Dolypeptide Human Dolypeptide Human Dolypeptide Sea urchin sulfatase pr Novel human diagno Sea urchin sulfatase pr Novel human diagno Human polypeptide C c clegans sulfata Novel human diagno Human polypeptide Novel human diagno Human sulfatase pr C c elegans sulfatase Novel human diagno Human sulfatase pr	AA.  AA.  tropic; neuroprotective; antibacterial; ntitumor; gene therapy; human.  Rudolph-Owen LA, Tsai F;  des useful for treating and diagnosing h as cerebrovascular diseases, acute degenerative diseases and tumor -
ABB61654 AAU33355 ABB52525 ABB52525 AAB85484 AAAM193920 AAX31920 AAX3123920 AAX31245 AAB51285 AAB51285 AAB51286 AAB51183 AAB51286 AAB51183 AAB51286 AAB51183 AAAM39414 AAM39414 AAM39414 AAM39414 AAM39414	ALIG 599 AA. 599 AA. nootrop ;; antit; t, Rudo t, Rudo such as such as such as
	(in; yy)  (in; yy)  (i)  (i)  (i)  (i)  (i)  (i)  (i)
996 4986 4986 4986 522 522 522 523 525 525 525 527 736 527 736 528 736 736 736 737 736 737 736 737 737 737	
222 222 223 244 244 245 255 265 265 265 265 265 265 265 265 26	Lan  1  23  23  23  23  23  23  1;  1;  1;  1;  1,  1,  1,  1,  1,  1,
6 0 3 5 5 6 0 3 5 5 6 0 3 5 5 6 0 3 5 5 6 0 3 5 5 6 0 3 5 5 6 0 3 5 5 6 0 3 5 5 6 0 3 5 5 6 0 3 5 5 6 0 3 5 5 6 0 3 5	1 85483 85483 852-2 SEP-2 SEP-2 SEP-2 Jan-2 Jan-2 Jan-2 Jan-2 Cksma ; 200 Cksma
111111111010202020202020202020202020202	710
	RESUL

~

```
Human; PRO; EST; expressed sequence tag; PCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98US-0084414
98US-0084441
                                                                                                                                                                                                                           20-MAR-1998;
20-MAR-1998;
25-MAR-1998;
26-MAR-1998;
27-MAR-1998;
                                                                                                                                                                                                                                                                                                                                          31-MAR-1998;
31-MAR-1998;
31-MAR-1998;
01-APR-1998;
01-APR-1998;
01-APR-1998;
01-APR-1998;
01-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                   08-APR-1998;
08-APR-1998;
09-APR-1998;
09-APR-1998;
                                                     Homo sapiens
                                                                       W09946281-A2
                                                                                                               08-MAR-1999;
                                                                                                                                                                                                        20-MAR-1998;
20-MAR-1998;
                                                                                                                                                                                                                                                                              27 - MAR-1998;
27 - MAR-1998;
                                                                                                                                                                                                                                                                                                   27-MAR-1998;
27-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        29-APR-1998;
29-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     06-MAY-1998
                                                                                                                                                                           12-MAR-1998,
                                                                                                                                                                                                                                                                                                                        30-MAR-1998
                                                                                                                                                                                                                                                                                                                                 30-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15-APR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    27-APR-1998
28-APR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              29-APR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21-APR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22-APR-1998
   ö
                                                 cerebrovascular diseases, infections such as acute meningitis, demyelinating diseases including multiple sclerosis, degenerative diseases affecting the cerebral correx including Alzheimer's disease and Pick disease, spinocerebellar degenerations including spinocerebellar ataxias including Friedreich ataxia, and ataxia telangiectesia, degenerative diseases affecting motor neurons including amyotrophic lateral sclerosis, inborn errors of metabolism such as leukodystrophics, toxic and acquired metabolic diseases, including vitamin deficiencies, and neurocutaneous syndromes (phakamatoses) including neurofibromotosis. The present sequence represents a human 23553 sulfatase polypeptide.
                    The invention provides 22438, 23553, 25278 or 26212 human sulfatase polypeptides and polynucleotides. The sulfatase genes and polypeptides are useful for treating disorders involving the brain such as
                                                                                                                                                                                                                                                                              ICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                MYTQRVQQILASHNPTKPIFLY1AYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCL 300
                                                                                                                                                                                                                                                                                                                                                                                                                                              DEAINNVTLALKTYGFYNNSIIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVH 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                       SPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILH 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                spllknkgtvckelvhltdwyptlislaegqidediqldgydiwetiseglrsprvdilh 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NIDPIYTKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGNPGYSDWVPPQSFSNLGPNRW 480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                       1 MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLA 60
                                                                                                                                                                                                                                                  FYRKECMPTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQ
                                                                                                                                                                                                                                                                                                                                                                                                               HNERITSSTGKSVWLFNITADPYERVDLSNRYPGIVKKLLRRLSOFNKTAVPVRYPFKDP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;
                                                                                                                                                                                              22; Length 599;
                                                                                                                                                                                                                   Indels
                                                                                                                                                                                              Score 2996; DB 22;
Pred. No. 5.3e-276;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY41701 standard; Protein; 515
Claim 9; Fig 15; 180pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Human PRO708 protein sequence.
                                                                                                                                                                                              99.58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                              Query Match 99.5
Best Local Similarity 99.6
Matches 548; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RSNPRLNGGV 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         599 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          07-DEC-1999
                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY41701;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ~
                                                                                                                                                                                                                                                           н
                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                                                                       241
                                                                                                                                                                                                                                                                                                                                                                                                                         241
                                                                                                                                                                                                                                                                                                                                                                                                                                             301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY41701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
  ò
                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
```

```
probe; blood coagulation disorder; cancer; cellular adhesion disorder; secreted protein; transmembrane protein.
primer; hybridisation;
                                                                                                                                                                                                                                                                                                  98US-0078936.
98US-0078939.
98US-0079294.
                                                                                                                                                                                                                                                                                                                                                             98US-0079663.
98US-0079664.
98US-0079689.
                                                                                                                                                                                                                                                                                                                                                                                                   9805 - 0079728
9805 - 0079728
9805 - 0079923
9805 - 0080105
9805 - 0080105
9805 - 0080105
9805 - 0080105
9805 - 0080327
9805 - 0080327
9805 - 0080333
9805 - 0080333
                                                                                                                                                                                                 98US-0077641.
98US-0077649.
98US-0077791.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98US-0081070.
98US-0081071.
98US-0081195.
98US-0081203.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98US-0081229.
98US-0081817.
98US-0081838.
98US-0081952.
98US-0082568.
98US-0082568.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98US-0082700.
98US-0082704.
98US-0082804.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98US-0082796.
98US-0083336.
98US-0083322.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98US-0083392.
98US-0083495.
98US-0083496.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0083500.
98US-0083545.
98US-0083554.
                                                                                                                                                                                                                                                                       98US-0078886
98US-0078910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0082767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98US-0083499
```

 $\sim$ 

us-09-495-823-7.rag

```
300
                                                                             300
                                                                                                       360
                                                                                                                                                          420
                                                                                                                                                                      protein; transmembrane protein; PRO; EST; cytostatic;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          , Eaton DL;
H, Gerritsen ME;
Hillan KJ;
ROY MA;
                                                                                                                    DEAINNVTLALKTYGFYNNSIIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVH
                 SPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILH
    FYRKECMPTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Wood
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gerber H,
                                                                                                                                                                                                            NIDPIYT---KAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGN 460
                                                                                                                                                                                                                           Human PRO708 (UNQ372) protein sequence SEQ ID NO:114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (P. Botstein D. Desnoyers L.,
E. Fong S. Gao W. Gerber H.,
5J. Grimaldi CJ. Gurney AL.
Napier MA, Pan J. Paoni NF.
PA. Tumas D. Williams PM, Wo.
                                                                                                                                                                                                                                                                                                                                                                                                                     tag; detection; cancer
                                                                                                                                                                                                                                                                                                      A
                                                                                                                                                                                                                                                                                                    AAB44257 standard; Protein; 515
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99WO-0S05028.
99US-0123957.
99US-0126773.
99US-0130232.
99US-0131445.
99US-0134287.
99US-0145698.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99WO-US31274.
2000WO-US00219.
2000WO-US00277.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99WO-US28565.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2000WO-US04341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99WO-US28313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99WO-US31243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2000WO-US00376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ashkenazi AJ, Baker KP, E
Ferrara N, Filvaroff E, F
Goddard A, Godowski PJ, G
Kljavin IJ, Kuo SS, Napie
Shelton DL, Stewart TA, T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99WO-US28551
                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GETH ) GENENTECH INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2000-611443/58.
                                                                                                                                                                                                                                                                                                                                                                                                                       sedneuce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2000-611443/
N-PSDB; AAC78483.
                                                                                                                                                                                                                                                                                                                                                                                                           secreted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200053756-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       26-JUL-1999;
29-OCT-1999;
30-NOV-1999;
02-DEC-1999;
16-DEC-1999;
30-DEC-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12-MAR-1999;
29-MAR-1999;
21-APR-1999;
28-APR-1999;
14-MAY-1999;
23-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                Homo sapiens
                                                                                                                                                                                                                                                                                                                                                      08-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 14-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                       expressed
                                                                                                                                                                                                                                                                                                                              AAB44257;
                                                                                                                                                                                                                                    421
    181
                            181
                                                       241
                                                                                                       301
                                                                                                                                 301
                                                                                                                                                          361
                                                                                                                                                                                    361
                                                                                                                                                                                                            421
                                                                                                                                                                                                                                                                                       AAB44257
                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                                                                                                                              g
                                                     Qγ
                                                                             8
                                                                                                       δ
                                                                                                                               g
                                                                                                                                                          ò
                                                                                                                                                                                 a
                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The present invention describes secreted and transmembrane polypeptides and their polynucleotides. The nucleotide sequences are useful as sources of probes, primers, for chromosome mapping, and for generation of antisense sequences. They can also be used to create transgenic animals. The proteins can be used to treat a variety of diseases and disorders, depending on their function. Diseases that may be treated include blood coagulation disorders, cancers and cellular adhesion adisorders. They may also be used to raise antibodies. AAX33891 to AAX34338, and AAX41685 to AAX41774 represent polynucleotide and polypertide sequence given in the exemplification of the present
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      secreted and transmembrane polypeptides and their polynucleotides, in for treating blood coagulation disorders, cancers and cellular
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QAGEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQP 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHWVCKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLA 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            20; Length 515;
                                                                                                                                                                                                                                                                                                                                                                                                        'n
                                                                                                                                                                                                                                                                                                                                                                                                        Chen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 2307.5; DB 20; Lengt
Pred. No. 1.5e-210;
3; Mismatches 15; Indels
                                                                                                                                                                                                                                                                                                                                                                                                         Baker KP,
                                                                                                                                                                                                                                                                                                                                                                                                         ر
ر
                                                                                                                                                                                                                                                                                                                                                                                                         Yuan
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Claim 12; Fig 43; 530pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                         Ą,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       э;
Э
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            76.68;
98US-0084598
98US-0084600
98US-0084637
98US-0084643
98US-0084643
98US-0084643
98US-008533
98US-0085538
98US-0085573
98US-0085580
98US-0085580
98US-0085607
98US-008507
98US-0086023
98US-0086023
98US-008603
98US-008603
98US-008603
98US-0086140
98US-0086430
98US-0086430
                                                                                                                                                                                                                                                                                                                                                                                                         Gurney
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                              (GETH ) GENENTECH INC.
                                                                                                                                                                                                                                                                                                                                                                                                       Goddard A,
                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1999-551358/46.
N-PSDB; AAZ33987.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               adhesion disorders
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        212
07 - MAY - 1998;
                                                                           07-MAY-1998;
13-MAY-1998;
13-MAY-1998;
13-MAY-1998;
15-MAY-1998;
15-MAY-1998;
                                                                                                                                                                                             15-MAY-1998;
15-MAY-1998;
                                                                                                                                                                                                                       15-MAY-1998;
18-MAY-1998;
                                                                                                                                                                                                                                                                                                                                         30-JUL-1998;
11-SEP-1998;
                                                                                                                                                                                                                                                22-MAY-1998
                                                                                                                                                                                                                                                            22-MAY-1998;
                                                                                                                                                                                                                                                                           MAY-1998;
                                                                                                                                                                     MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 434;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                                                                                                                                                                                                                                      22-MAY-19
28-MAY-19
                                                                                                                                                                                                                                                                                                                                                                                                        Wood WI,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local 9
                                                                                                                                                                                    .5-MAY-1
                                                                                                                                                                                                                                                                                                               28-MAY-1
                                                                                                                                                                                                                                                                                                                             28-MAY-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     useful
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       New
```

Novel PRO polypeptides and polynucleotides used in detection methods,

g

ð g

g ò

ð

us-09-495-823-7.rag

28-FEB-2001; 2001WO-US06520

WO200168848-A2

20-SEP-2001

```
4,
                                                                            AAC78458 to AAC78599 represent polynucleotide and EST (expressed sequence tag) sequences which encode secreted or transmembrane PRO polypeptides. The PRO polynucleotides and polypeptides have cytostatic activity. The polynucleotides and polypeptides can be used for detecting the presence of PRO polypeptides in samples, for linking bloactive molecules to cells and for modulating biological activities of cells, using the polypeptides for specific targeting. The polypeptide targeting the polypeptide for specific targeting. The polypeptide targeting the polypeptide pairs provide specific targeting of bioactive molecules to cells. AAC78600 to AAC7887 represent PCR primers and probes used in the isolation of the PRO polynucleotide sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha; blood; chondrocyte cell; cell proliferation; cell differentiation; colon; adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121 ICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FYRKECMPTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQ 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRSIININRRYAAMLSCL 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DEAINNVTLALKTYGFYNNSIIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVH 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILH 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QAGEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQP 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              dealnnvtlalktygfynnsiilyssdnggaptaggsnwplrgskgtyweggiravgfvh 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
                                                                                                                                                                                                                                                                                                                                                                              15; Indels 11; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                1 MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLA 60
                                                                                                                                                                                                                                                                                                                                                                                                                               target bioactive molecules to specific cells, and to modulate
                                                                                                                                                                                                                                                                                                                                              21; Length 515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NIDPIYT---KAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGN 460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  nidp-ytprqkmapgqqamgsg----tlqss---qpsecstgn 455
                                                                                                                                                                                                                                                                                                                                           Score 2307.5; DB 2;
Pred. No. 1.5e-210;
3; Mismatches 15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human PRO polypeptide sequence #38.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ¥.
                                                  Claim 12; Fig 43; 636pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAU29061 standard; Protein; 515
                                                                                                                                                                                                                                                                                                                                            76.6%;
93.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                           Query Match 76.6
Best Local Similarity 93.7
Matches 434; Conservative
                cellular activities
                                                                                                                                                                                                                                                                                          515 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAU29061;
                                                                                                                                                                                                                                                                                              Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
AAU29061
 S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
```

```
Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention. The PRO polypeptides and their associated nucleic acids can be used to detect the presence of a tumour in a mammal by comparing the level of expression of a PRO polypeptide in a test sample of cells from the animal and a control sample of normal cells, whereby a higher level of expression in the test sample indicates the presence of a tumour in the mammal. Mammals include dogs, cattle, horses, sheep, pigs, goats and rabbits but are preferably human. The polypeptides can be used to stimulate tumour necrosis factor (TMF) alpha release from human blood, when contacted with it. A specific polypeptide can be used to stimulate the proliferation or differentiation of chondrocyte cells. The PRO proteins can be used to determine the presence of tumours and also
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Novel nucleic acids encoding PRO polypeptides, used to diagnose the presence of tumours, such as prostate and breast tumours, in mammals and to screen for modulators of the compounds -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Godowski PJ, Gurney AL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Goddard A, Goo
--- wt, Zhang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Wood WI,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ļ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 11; Fig 76; 774pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Watanabe CK,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Chen J, Desnoyers
                                                                                                                                                                                                                                                                                                                                                     2000US-191007P.
2000US-191048P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2000US-193053P.
2000WO-US08439.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2000US-194449P.
2000US-194647P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2000US-195975P.
2000US-196000P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2000US-196187P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000US-198121P.
2000US-198585P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2000US-199397P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2000WO-US14042.
                                                                                                                                                                                                                                                                                                                                 2000US-190828P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2000US-196690P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2000US-196820P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000US-199550P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2000US-199654P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2000WO-US14941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2000US-0644848
2000WO-US23328
                                                                                                                                                                                                                                                         2000US-189320P
                                                                                                                                                                                                                                                                                                      2000WO-US06884
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000US-201516P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2000WO-US13705
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2000WO-US15264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2000US-209832P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GETH ) GENENTECH INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2001-602746/68.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Smith V,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-PSDB; AAS45962
                                                                                                                                                                              02-MAR-2000; 203-MAR-2000; 204-MAR-2000; 2000; 214-MAR-2000; 211-MAR-2000; 211-MAR-2000; 211-MAR-2000; 228-MAR-2000; 229-MAR-2000; 229-MAR-2000; 204-APR-2000; 211-APR-2000; 211-APR-200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      18-APR-2000; 25-APR-2000; 25-APR-2000; 25-APR-2000; 25-APR-2000; 25-APR-2000; 217-MAY-2000; 217-MAY-2000; 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22-MAY-2000; 2
30-MAY-2000; 2
02-JUN-2000; 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   05-JUN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Baker KP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pan J,
```

S

N-PSDB; AAH46862, AAH46866.

```
4;
susceptibility to tumour development, particularly adrenal, lung, colon, breast, prostate, rectal, cervical, or liver tumours, in mammalian subjects. The oligonuclectide probes specific for the PRO nucleic acids can be used for genetic analysis of individuals with genetic disorders.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sulfatase; 25278 sulfatase; nootropic; neuroprotective; antibacterial; antiinflammatory; vasotropic; antitumor; gene therapy; human.
                                                                                                                                                                                                                                               300
                                                                                                                                                                                        QAGEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQP 120
                                                                                                                                                                                                180
                                                                                                                                                                                                                                                                                                                                MYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCL 300
                                                                                                                                                                                                                                                                                                                                                                              DEAINNVTLALKTYGFYNNSIIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVH 360
                                                                                                                                                                                                                                                                                                                                                                                           SPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILH 420
                                                                                                                                          9
                                                                                                                                                                9
                                                                                                                                                     MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLA
                                                                                                                                                                                                                                     ICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLG
                                                                                                                                                                                                                                                                                 FYRKECMPTRRGFDTFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQ
                                                                                             Length 515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Tsai F;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               nidp-ytprqkmapgqqamgsg----tlqss---qpsecstgn 455
                                                                                             DB 22;
                                                                                          Score 2307.5; DB 22
Pred. No. 1.5e-210;
3; Mismatches 15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NIDPIYT - - - KAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Rudolph-Owen LA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ä.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sulfatase polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB85482 standard; Protein; 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Glucksmann MA, Williamson M,
                                                                                          76.6%;
ilarity 93.7%;
Conservative 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (MILL-) MILLENNIUM PHARM INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            31-JAN-2001; 2001WO-US03266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31-JAN-2000; 2000US-0495823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2001-476214/51.
                                                                                                        Similarity
                                                         515 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200155411-A2.
                                                                                          Query Match
Best Local Simi
Matches 434;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Human 25278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    02-AUG-2001
                                                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAB85482;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             421
                                                                                                                                                                                                                                                                                                                                                                                                  301
                                                                                                                                                                                                                                                                                                                                                                                                                                                 361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ношо
                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                 181
                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                241
                                                                                                                                                                                                                                                                                                                                                                              301
                                                                                                                                                                                                                                                                                                                                                                                                                           361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       421
                                                                                                                                                                                                                                    121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8×66666
                                                                                                                                         ò
                                                                                                                                                              g
                                                                                                                                                                                       ö
                                                                                                                                                                                                           g
                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ď
```

```
The invention provides 22438, 23553, 25278 or 26212 human sulfatase polypeptides and polypucleotides. The sulfatase genes and polypeptides are useful for treating disorders involving the brain such as cerebrovascular diseases, infections such as acute meningitis, demyelinating diseases including multiple sclerosis, degenerative diseases affecting the cerebral cortex including Alzheimer's disease and Pick disease, spinocerebellar degenerations including spinocerebellar ataxias including Friedreich ataxia, and ataxia telangiectasia, degenerative diseases affecting motor neurons including amyotrophic lateral sclerosis, inborn errors of metabolism such as leukodystrophies, toxic and acquired metabolic diseases, including vitamin deficiencies, and neurocutaneous syndromes (phakamatoses) including neurofibromocosis, and neurocutaneous syndromes (phakamatoses) including neurofibromocosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         352
                             Novel human sulfatase polypeptides useful for treating and diagnosing sulfatase-related disorders such as cerebrovascular diseases, acute meningitis, multiple sclerosis, degenerative diseases and tumor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               83 ADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITGKYQIHTGLQH 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27 MGALAGFWILCLLTYGYLSWGQALEEEEEGALLA----QAGEKLEPSTTSTSQPHLIFIL 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDTFFGSLLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       323 IYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWYP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          383 TLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPIYTKAKNGSWAAGYGIWN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          443 TAIQSAIRVQHWKLLTGNPGYSDWVPPQSFSNLGPNRWHNERITSSTGKSVWLFNITADP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SGDYYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQILASHNPTKPIFLY
                                                                                                                                                                                                                                                                                                                                                                                                                  Length 569;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    81; Mismatches 103; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             YERVDLSNRYPGIVKKLLRRLSQFNKTAVPVRYPPKDPRSNPRLNGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. 8.1e-168;
                                                                                                                                                                                                                                                                                                                                                                                                                  61.7%; Score 1859.5; 62.6%; Pred. No. 8.1e
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ä
                                                                                               Claim 9; Fig 10; 180pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAU14421 standard; Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 62.6
Matches 330; Conservative
                                                                                                                                                                                                                                                                                                                                                                   569 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAU14421;
                                                                                                                                                                                                                                                                                                                                                                     Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAU14421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 XX
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

5;

IIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDTFFGSLLGS

144

```
The invention relates to polynucleotides encoding novel human proteins or their active domains. The polypeptides, polynucleotides and antibodies raised against the polypeptides are used in a method of treatment of a mammal and prevention of disorders caused by the aberrant protein expression or activity. The polypeptides can be used as molecular weight markers, food supplements, and in antibody production. The polypeptides are used to identify compounds which bind to the polypeptides. Polynucleotides of the invention are used as probes and production of recombinant proteins, and in generating anti-sense DNA or RNA and in gene therapy. Polypeptides of the invention can be used to target drugs to a tumour, in assays to determine biological activity, to raise antibodies/elicit an immune response, to determine quantitative protein levels, as tissue markers, and to isolate receptors or ligands. Polypeptides of the invention may also be useful in treating platelet illoments and/or nerve tissue, wound healing, treating burns, promoting the proliferation, differentiation and survival of stem cells, as a
                                                                                   immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic; anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral; antibacterial; antialteratic; dermatological; haemostatic; antiasthmatic; thrombolytic; immunogen; antibody; gene therapy; neurological disorder; Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis; tissue regeneration; immune disorder.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            contraceptive, treating osteoporosis and osteoarthritis, anaemia,
Alzheimer's, Parkinson's and Huntington's diseases, amylotrophic lateral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sclerosis, stroke, immune deficiencies resulting from bacterial, viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Isolated polypeptides useful for treating anti-inflammatory diseases, nervous system disorders, and for regenerating bone and cartilage -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        fungal infection or from autoimmunity, cancer, allergy, asthma, graftversus.host disease, eczema, haemophilia, thrombosis, anti-inflammatory diseases, nervous system disorders, and infection. The present sequence represents a protein of the invention.
                                                                novel protein; Antianaemic; osteopathic; antiinflammatory;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Example 4; Page 809-810; 894pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Drmanac RT;
                                                                                                                                                                                                                                                                                                                                                                               25-JAN-2001; 2001WO-US02623
                                                                                                                                                                                                                                                                                                                                                                                                                         25-JAN-2000; 2000US-0491404.
                    novel protein #292.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2001-451939/48.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       289 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (HYSE-) HYSEQ INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Liu C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAS22726
                                                                                                                                                                                                                                                                                       WO200155437-A2.
                                                                                                                                                                                                                                              Homo sapiens.
                                                                                                                                                                                                                                                                                                                                   02-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Tang YT,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                Human;
                  Human
```

ó Gaps 24 MLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLAQAGEKLEPSTTSTSQPHLIFILA 83 ö Length 289; Indels Pred. No. 5,8e-140; ; 0 DB 22; Mismatches 51.9%; Score 1562; 100.0%; Pred. No. 5 ; 0 Matches 289; Conservative Similarity Query Match Local à 셤

DDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITGKYQIHTGLQHS 143 

84 61

ð

æ

Gaps

57;

Conservative

Matches 278;

Similarity

The present invention describes a recombinant human iduronate 2-sulfatase (IDS). The recombinant IDS is more highly glycosylated than the naturally occurring enzyme isolated from human tissue. The recombinant human IDS can be produced in Chinese Hamster Overy (CHO) cells or in a human cell. The recombinant IDS comprises a fusion protein. It is a morpolysaccharidosis inhibitor and can be used in gene therapy. The recombinant IDS is useful in treating and diagnosing subjects suffering from or suspected of having IDS deficiency disorders, e.g. thunter syndrome (mucopolysaccharidosis type II). The present sequence represents a human sulfatase which is used in comparison with the IDS sequence in an example from the present invention. New highly glycosylated recombinant human iduronate 2-sulfatase (IDS) useful for diagnosing or treating subjects suspected of having or suffering from IDS deficiency disorders, e.g. Hunter syndrome GDYYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQILASHNPTKPIFLYI Occhiodoro T, Bielicki J, Clements PR; DB 22; Length 533; Human; iduronate 2-sulfatase; IDS; glycosylated; gene therapy; mucopolysaccharidosis inhibitor; IDS deficiency disorder; Hunter syndrome; mucopolysaccharidosis type II. 47.0%; Score 1415.5; DB 22; Lengt 49.1%; Pred. No. 1.4e-125; iive 80; Mismatches 151; Indels 264 AYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLALK Human sulfatase protein B SEQ ID NO:13. (WOME-) WOMEN'S & CHILDREN'S HOSPITAL. AAB51184 standard; Protein; 533 AA Example 1; Fig 3; 53pp; English. 92US-0991973. 94US-0345212. 91US-0790362. 99US-0249003. (mucopolysaccharidosis-II) (first entry) Anson DS, (Morris CP; WPI; 2001-060076/07. 533 AA; Homo sapiens 12-FEB-1999; 21-MAR-2001 28-NOV-1994; 12-NOV-1991; US6153188-A. 28-NOV-2000. 17-DEC-1992; Hopwood JJ, Wilson PJ, AAB51184; Sequence Query Match Best Local 204 AAB51184 RESULT g g ð or

ò 9 ò 8 ò g ò g ò g ò 셤 ò 셤 ò g ò 合 ò 셤

```
WPI; 2001-451939/48.
                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                              366 AA;
           N-PSDB; AAS22490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16-0CT-2001
                                                                                                                                                                                                                                                                                                                                                                   Best_Local Sir
Matches 199;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAM25415;
                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local S
                              Isolated
                                        nervous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   342
                                                                                                                                                                                                                                                                                                                                                                                              16
                                                                                                                                                                                                                                                                                                                                                                                                                                     97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAM25415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Human; novel protein; Antianaemic; osteopathic; antiinflammatory; immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic; anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral; antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic; thrombolytic; immunogen; antibody; gene therapy; neurological disorder; Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis; tissue regeneration; immune disorder.
                                                                          180
                                                                                                                                                                                                     415
                                                                                                                                                                                                                                           120
                                                                                                                                                                                                                                                                      464
                   33
                                                                                                                          61 QAGEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQP
                                               ICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMYGKWHLG
                                                                                    FYRKECMPTRRGFDTFFGSLLGSGDYYTHYKCD----SPGMCGYDLYENDNAAWDYDNG
                                                                                                                                                     IYSTQMYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRSIININRRYAA
                                                                                                                                                                                          MLSCLDEAINNVTLALKTYGFYNNSIIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRA
                                                                                                                                                                                                                                                                     VDILHNIDPIY-----TKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGNPGYS
                                                                                                                                                                                                                                                                                                          DWVPPOSFSNLGPNRWHNERITSS--TGKSVWLFNITADPYERVDLSNRYPGIVKKLLRR
                                                                                                                                                                                                                                                                                                                      1 MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLA
                                                                                                                                                                                                                               VGFVHSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPR
                  mgprgaaslprgpgprrlllpvvlp------11111-
                                                                                                                                                                                                                                                                                                                                               LSQFNKTAVPVRYPPKDPRSNPRLNG 548
                                                                                                                                                                                                                                                                                                                                                                                                                 ¥.
                                                                                                                                                                                                                                                                                                                                                                502 lqfyhkhsvpvyfpaqdprcdpkatg
                                                                                                                                                                                                                                                                                                                                                                                                               AAU14185 standard; Protein; 366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       rang YT, Liu C, Drmanac RT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25-JAN-2001; 2001WO-US02623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25-JAN-2000; 2000US-0491404
                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protein #56.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HYSE-) HYSEQ INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO200155437-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                      24-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Human novel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            32-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                   AAU14185;
                                                                                                                                                                                                                                                                                                                                                                                              æ
                                                                          121
                                                                                                                                                     236
                                                                                                                                                                       209
                                                                                                                                                                                          296
                                                                                                                                                                                                             269
                                                                                                                                                                                                                                                  329
                                                                                                                                                                                                                                                                     416
                                                                                                                                                                                                                                                                                       389
                                                                                                                                                                                                                                                                                                           465
                                                                                                                                                                                                                                                                                                                            449
                                                                                                                                                                                                                                                                                                                                               523
                                                                                                              181
                                                                                                                                                                                                                                356
                                                                                                                                                                                                                                                                                                                                                                                              RESULT
```

```
The invention relates to polynucleotides encoding novel human proteins or their active domains. The polypeptides, polynucleotides and antibodies raised against the polypeptides are used in a method of treatment of a mammal and prevention of disorders caused by the aberrant protein expression or activity. The polypeptides can be used as molyneular weight markers, food supplements, and in antibody production. The polypeptides are used to identify compounds which bind to the polypeptides. Polynucleotides of the invention are used as probes and production of recombinant proteins, and in generating anti-sense DNA or RNA and in gene therapy. Polypeptides of the invention can be used to target drugs to a tumour, in assays to determine biological activity, to raise antibodies/elicit an immune response, to determine duantitative protein levels, as tissue markers, and to isolate receptors or ligands. Polypeptides of the invention may also be useful in treating platelet disorders, stem cell disorders, regenerating bone, cartilage, tendon, ligament and/or nerve tissue, wound healing, treating burns, promoting the proliferation, differentiation and survival of stem cells, as a contraction and survival of stem cells.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          contraceptive, treating estepoporosis and esteoarthritis, anaemia, contraceptive, treating esteoporosis and esteoarthritis, analycrophic lateral Alzheimer's, Parkinson's and Huntington's diseases, amylotrophic lateral sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Human; cancer; ulcer; HIV infection; human immunodeficiency virus; antiinflammatory; antiirheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia; antiaggregant; haemostatic; vulnerary; antiulcer; osteopathic; eczema;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITGKYQ 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            136 IHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDT 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
J polypeptides useful for treating anti-inflammatory diseases, system disorders, and for regenerating bone and cartilage -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sclerosis, stroke, immune deficiencies resulting from bacterial, virifungal infection or from autoimmunity, cancer, allergy, asthma, graft-versus-host disease, eczema, hemophilia, thrombosis, anti-inflammatory diseases, nervous system disorders, and infection. The present sequence represents a protein of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         282 ihtglqhsiirptqpnclpldnatlpqklkevgysthmvgkwhlgfyrkecmptrrgfdt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLAQAGEKLEPSTTSTSQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 366;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 1079; DB 22;
Pred. No. 8.3e-94;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0; Mismatches
                                                                                                                                  Example 4; Page 557-558; 894pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human protein sequence SEQ ID NO:930.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAM25415 standard; Protein; 187 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    196 FFGSLLGSGDYYTHYKCDSP 215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     35.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
```

ö

œ

AAM2563. The proteins can have activities based on the tissues and cells they are expressed in. such as: antinflammatory; antinfheumatic; cells they are expressed in. such as: antinflammatory; antinfheumatic; cells they are expressed in. such as: antinflammatory; antinfheumatic; cardiant; cardiants. inflammanic artiangenic; antidactic; vulnerary; cardiances; osteopathic; dermatological; antiallergic; antiasthmatic; antidabetic; cytostatic; neuroprotective; antidepressant; nootropic; antidation; hem can be used in gone therapy, antisense therapy and vaccine production, The proteins and polynucleotides are useful for screening for agonists of antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, anaemia, platelet disorders, thromboorytopaenia, wounds, burns, ulcers, costeoporosis, severe combined immunodefficiency, eczema, allergic neuroprotective; antidepressant; nootropic; antiparkinsonian; infection; immunostimulant; gene therapy; antisense therapy; vaccine; inflammation; antianaphylactic; rheumatold arthritis; septic shock; pancreatitis; cardiac dysfunction; neuropathylogy; cardiac anaphylaxis; autoimmuity; genetic disease; haematopoietic disorder; platelet disorder; asthma; thrombocytopaenia; osteoporosis; severe combined immunodeficiency; allergic rhinitis; diabetes; multiple sclerosis; depression; Alzheimer's disease; Parkinson's disease; neurodegenerative disorder; rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzhelmer's disease, Parkinson's disease, neurodegenerative and polynucleotides encoding polypeptides, useful for Hagnosis of e.g. cancer, ulcers and HIV infection dermatological; antiallergic; antiasthmatic; antidiabetic; Claim 20; Page 200; 1217pp; English. treatment and diagnosis of e.g. 99US-0471275. 2000US-0488725. 2000US-055317. Liu C, Drmanac RT; 22-DEC-2000; 2000WO-US35017 neurological disorders. neurological disorder 2001-457603/49. 187 AA; (HYSE-) HYSEQ INC N-PSDB; AAH99356 WO200153455-A2. Isolated human 21-JAN-2000; 25-APR-2000; 23-DEC-1999; 26-JUL-2001 Alzheimer's Tang YT, Sequence Ношо WAS COULDED COULD COULD

KYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRG 192 FDTFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQILAS 252 67 ö Indels ö Score 989; DB 22; Pred. No. 1e-85; 1; Mismatches Query Match 32.8%; Best Local Similarity 99.4%; Matches 179; Conservative 133 œ 68 193 ò ò 8 g

ö

15;

213 DSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQILASHNPTK-PIFLYIAYQAVHS-

ò

HNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLALK 312

253

ò

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention sequences (ABLIG176-ABL30511), expressed DNA sequences (ABLIG175) and the encoded proteins The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences. New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell 212 HGS-EIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITGKYQIHTGLQHSIIRPTQPNC 152 Gaps 34 WILCLLTYGYLSWGQALEEEEEGALLAQAGEKLEPSTTSTSQPHLIFILADDQGFRDVGY LPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDTFFGSLLGSGDYYTHYKC Drosophila; developmental biology; cell signalling; insecticide; 98; Disclosure; SEQ ID NO 18678; 21pp + Sequence Listing; English. Length 542; Indels Drosophila melanogaster polypeptide SEQ ID NO 18678 DB 22; Mismatches 30.5%; Score 920; Pred. No. Myers AA 542 79; PWD, ABB63962 standard; Protein; 23-MAR-2000; 2000US-191637P. 11-JUL-2000; 2000US-0614150. 23-MAR-2001; 2001WO-US09231 (first entry) Ľį Conservative Drosophila melanogaster. Adams M, WPI; 2001-656860/75. (PEKE ) PE CORP NY. Similarity 542 AA; N-PSDB; ABL08065 WO200171042-A2. pharmaceutical interactions 26-MAR-2002 Mac Local S... 213; 27-SEP-2001 Venter JC, ABB63962; Seguence Query Match Matches 128 94 153 **ABB63962** g g δ 8 ŏ ò a

14;

Gaps

57;

249

304 263 359

```
The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            419 LHNIDPIY------TKAKNGSWAAG-YGIW------NTAIQSAIRVQHWKL 456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             457 LTGNPGY-----SDWVPPQSFSNLGPNRWHNERITSSTGKSVWLFNITADPYE 504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    442 llgnrgltkdrirgmrseatetcppiegqnpleshfkcepl-----kapcffdlakdpce 496
                                                                                                                                                                                                                                    73 TSQPHLIFILADDQGFRDVGYHGS-EIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFIT 131
                                                                                                                                                                                                                                                                                                                     GKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRR 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        :| :| ::
382 ihvldevfgyssymrdtlkyvngssfkgrydqwlgeletneddplgesyeqhvlasdvqs
                                                                                                                                                                                                                                                                                                                                           305 NNVTLALKTYGFYNNSIIIYSSDNGGQPTAG----GSNWPLRGSKGTYWEGGIRAVGFV
                                                                                                                                                                                                                                                                                                                                                                                                                                HSPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRV-DI
                                                                                                                                                                                                                                                                                                                                                                                                        GFDTFFGSLLGSGDYYTHY - - KCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical.
                                                                                                                                                    Length 579;
                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        505 RVDLSNRYPGIVKKLLRRLSQFNKTAVP-VRYPPKDPRSNPRLNGG 549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Drosophila melanogaster polypeptide SEQ ID NO 11754
                                                                                                                                                  28.1%; Score 846.5; DB 22; 36.5%; Pred. No. 2.4e-71;
                                                                                                                                                                       ; Pred. No. 2.4e-71;
85; Mismatches 192;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Protein; 996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2000US-191637P.
2000US-0614150.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        23-MAR-2001; 2001WO-US09231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Drosophila melanogaster
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ABB61654 standard;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (PEKE ) PE CORP NY
                                                                                                                                                                         Similarity
                                                                                         AA;
                                                                                       579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO200171042-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               23-MAR-2000;
11-JUL-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               27-SEP-2001
                                                                                                                                                                         Best Local Sim
Matches 192;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ABB61654;
                                                                                         Sequence
                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           250
                                                                                                                                                                                                                                                                            25
                                                                                                                                                                                                                                                                                                                     132
                                                                                                                                                                                                                                                                                                                                                              85
                                                                                                                                                                                                                                                                                                                                                                                                        192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ABB61654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
      2255×8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  THE STATE OF THE S
                                                                                                                                                                                                                                                                                                                     ó
                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                       ŏλ
                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from prosophila. The invention useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABLi6176-ABL30511), expressed DNA sequences (ABL16175) and the encoded proteins
                                                                                                                                                                                                                                    430
                                                                                                                                                                                                                                                                                                                                                                                                        526
                                                                                                        261
                                                                                                                                                  381
                                                                                                                                                                                           320
                                                                                                                                                                                                                                                                            380
                                                                                                                                                                                                                                                                                                                     484
                                                                                                                                                                                                                                                                                                                                                                                                                              DNGGQPTAG----GSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKEPVHITDWY
                                                                                                                                                                         ||||: |||| : |:|| || : |||||| |: ||||| ||: ||||| ||: ||||| ptlleaagggpalsnlskqidggsiwralvgdkasprlnvlhniddiwgsaalsvgdwkl
                                                                                                                                                                                                                                                                                                                                             I-----TSSTGKSV-----WLFNITADPYERVDLSNRYPGIVKKLLRRLSQF
                                                                 ----PLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYSS
                                                                                                                                                                                                                                PTLISLAEGQ---IDEDIQLDGYDIWETISEGLRSPRVDILHNIDPIYTKAK-----
                                                                                                                                                                                                                                                                                                                 ----NGSWAAGYGIWNTAIQSAIRVQHWKLLTGNPGYSDWVPPQSFSNLGPNRWHNER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; SEQ ID NO 16530; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ID NO 16530
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Œ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Drosophila melanogaster polypeptide SEQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Myers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABB63246 standard; Protein; 579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PWD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23-MAR-2000; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23-MAR-2001; 2001WO-US09231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | ||||
| natavppsnkpadpradpr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŗ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NKTAVPVRYPPKDPRSNPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Drosophila melanogaster
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Adams M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2001-656860/75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (PEKE ) PE CORP NY
                          ----qvaydlh-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-PSDB; ABL07349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               pharmaceutical.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200171042-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                interactions -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Drosophila;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       27-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Venter JC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ABB63246;
                                                                                                                                                                                                                                                                                                                                                                                                   485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             494
                                                                                                                                               327
                                                                                                                                                                                                                                  382
                                                                                                                                                                                                                                                                                                                   431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      527
                       157
                                                                                                                                                                                                                                                                         321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABB63246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                             ò
                                                                                                                                                                                    g
                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                      ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

441

AA.

AAU23335 standard; Protein; 128

```
AAU23335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                18;
                                                                                                                                                                                                                                            ïs
                                                                                                                                                                                                                   The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention i useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176 ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABS5737-ABB20022). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                            New isolated nucleic acid detection reagent for detecting 1000~\mathrm{or} more genes from Drosophila and for elucidating cell signalling and cell-cell interactions .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FRDVGYHGS-EIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITGKYQIHTGLQHSIIR 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    147 PTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDTFFGSLLGSGDY 206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PIFLYIAYQAVHS-----PLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLALK 312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         272 stdqlensivifysdnga-psvgmfsntgsnfplrgqkntpweggvrvagaiwssglgar 330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRV--DILHNIDPI 425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------dvwrl---salqmgqwkyvngttasgrydsvltyrelddldprdsry 432
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             207 Y-----THYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQILASH-NPTK 257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYGFYNNSIIIYSSDNGGQPTAG----GSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNK 367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ----SVWLFNITADDY 503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25.0%; Score 752; DB 22; Length 996;
31.4%; Pred. No. 5.8e-62;
ive 98; Mismatches 188; Indels 116; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | | :||| | :||| | 3 glsafllllcl------grvksdesaa------arrpniiimaddmg 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 fddvsfrggrefltpnidalayhgrlldrlyapamctpsrgallsgrypihtgtqhfvis 98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GALAGFWILCLLTYGYLSWGQALEEEEEGALLAQAGEKLEPSTTSTSQPHLIFILADDQG 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    : :| |: | : :: | : : | farrskmpvanys-----lgydfrrnmelec-rdrgvyvtdlltaeaerlikdhadkeq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   YTKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTGNPG----YSDWVPPQSFSNLGP----
                                                                                                                                                                              Disclosure; SEQ ID NO 11754; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             eqnnlvyserhsdvltalrrrvqelrasasrpgnrasmpeadptlh 538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     504 ERVDL--SNRYPGIVKKLLRRLSQFNKTAVPVRYPPKDPRSNPRLN 547
                   ΕK
                   Myers
                 Li PWD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 31.4%
Matches 184; Conservative
                 Adams M,
                                                  WPI; 2001-656860/75
                                                                                                                                                                                                                                                                                                                                                                                                                                                    996 AA;
                                                                      N-PSDB; ABL05757
               Venter JC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                313
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       493
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           478
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
```

13

RESULT

```
Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase; ligase; hyperproliferative disorder; immunodeficiency disorder; autoimmune disorder; neurological disorder; metabolic disorder; inflammatory disorder; cardiovascular disorder; reproductive disorder; blood-related disorder; infectious disorder; anti arthritic; nephrotropic; anticoagulant.
                                                                                                                 Novel human enzyme polypeptide #421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24-FEB-2000; 2000US-0184664.
16-MAR-2000; 2000US-0184664.
16-MAR-2000; 2000US-01846550.
16-MAR-2000; 2000US-0184123.
18-MAY-2000; 2000US-0198123.
19-MAY-2000; 2000US-0198123.
19-MAY-2000; 2000US-0198123.
11-MAY-2000; 2000US-0216880.
11-JUL-2000; 2000US-0216880.
11-JUL-2000; 2000US-0216880.
11-JUL-2000; 2000US-0216880.
11-JUL-2000; 2000US-0216880.
11-JUL-2000; 2000US-021880.
14-AUG-2000; 2000US-0222213.
14-AUG-2000; 2000US-0222214.
14-AUG-2000; 2000US-0222214.
14-AUG-2000; 2000US-0222214.
14-AUG-2000; 2000US-0222214.
14-AUG-2000; 2000US-0222217.
14-AUG-2000; 2000US-022217.
14-AUG-2000; 2000US-022217.
14-AUG-2000; 2000US-022217.
14-AUG-2000; 2000US-022217.
14-AUG-2000; 2000US-022217.
14-AUG-2000; 2000US-022217.
15-AUG-2000; 2000US-022217.
16-SEP-2000; 2000US-022217.
17-SEP-2000; 2000US-022217.
18-AUG-2000; 2000US-0222217.
18-AUG-2000; 2000US-0222217.
1
                                                                                                                                                                                                                                                                                                                                                                                                                                            2000US-0179065.
2000US-0180628.
2000US-0184664.
2000US-0186350.
2000US-0189874.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2000US-0231413
2000US-0231414
                                                                                                                                                                                                                                                                                                                                                                                                         17-JAN-2001; 2001WO-US01239
                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                WO200155301-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         08-SEP-2000;
                                                                                                                                                                                                                                                                                          Homo sapiens
                                                                         17-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                       02-AUG-2001
                                  AAU23335;
```

us-09-495-823-7.rag

Page 11

14-SEP-2000; 14-SEP-2000; 14-SEP-2000; 21-SEP-2000;

14-SEP-2000;

-SEP-2000;

-SEP-2000; -SEP-2000;

29-SEP-2000;

-SEP-2000

27-SEP-2000;

29-SEP-2000

```
17-NOV-2000, 2
01-DEC-2000, 2
05-DEC-2000, 2
05-DEC-2000, 2
05-DEC-2000, 2
06-DEC-2000, 2
06-DEC-2000, 2
08-DEC-2000, 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 26-MAR-2002
                                                                                                                                                                                                                                                                                                                                invention.
                                                                                                                             Rosen CA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ABB63259;
                                                                                                                                                                                       diseases
                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ABB63259
                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AX SX EX
  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                          20000S-0246474.
20000S-0246475.
20000S-0246477.
20000S-0246477.
20000S-0246523.
20000S-0246524.
20000S-0246525.
20000S-0246526.
20000S-0246526.
                                                                                                                                                                                                                                                                                                                                                                                               2000US-0246611.
2000US-0246613.
2000US-0249207.
                                                                                                                                                                                                                                                                     2000US-0241809.
2000US-0241826.
                                                                                                                                                  2000US-0236368
                                                                                                                                                                 0000US-0236370
                                                                                                                                                                                                                                                              2000US-0241808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2000US-0249299
                2000US-0231968
                                     2000US-0232399
                                                                                               2000US-0234997
                                                                                                                                          2000US-0236367
                                                                                                                                                                                                                                                                                   2000US-0244617
                                                                                                                                                                                                                                                       2000US-0241787
```

20-OCT-2000; 20-OCT-2000;

13-OCT-2000 13-OCT-2000

31-NOV-2000;

NOV-2000; NOV-2000; NOV-2000; NOV-2000; NOV-2000;

NOV-2000

NOV-2000

-NOV-2000; -NOV-2000; -NOV-2000; -NOV-2000;

-NOV-2000;

NOV-2000;

17-NOV-2000;

NOV-2000;

polypeptides, and the CDNA (ASA0785-SAS41684) and genomic sequences canceding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of a wide range of diagnosis, treatment, prevention and/or prognosis of a wide range of immunodeficiency disorders (e.g. cancer), immunodeficiency disorders (e.g. AlbS) autoimmune disorders (e.g. arthritis), neurological disorders (e.g. atheometry disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. thempolytia), reproductive disorders (e.g. inflammatory disorders (e.g. inflammatory disorders (e.g. inflammatory disorders (e.g. inflammatory alsorders (e.g. inflammatory alsorders (e.g. inflammatory alsorders (e.g. inflammatory) and alsorders (e.g. inflammatory) and alsorders (e.g. inflammatory) and alsorders (e.g. inflammatory) and alsorders (e.g. inflammatory). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences. preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous 90 DVGYHGSEIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITGKYQIHTGLQHSIIRPTQ 149 150 PNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDTFFGSLLGS-GDYYT 208 Gaps The present invention relates to the isolation of novel human enzyme Length 128; Novel polypeptides and polynucleotides useful for diagnosing, 20.0%; Score 603.5; DB 22; Length 94.2%; Pred. No. 2.9e-49; ive 0; Mismatches 6; Indels Claim 11; SEQ ID No 1331; 1180pp; English ABB63259 standard; Protein; 486 Ruben SM; 2000US-0251030. 2000US-0251988. 2000US-0256719. 2000US-0251479. 2000US-0251856. 2000US-0251868. 2000US-0251869. (HUMA-) HUMAN GENOME SCI INC 2000US-0254097. 2001US-0259678. 2000US-0251990 2000US-0250160 2000US-0250391 2000US-0251989 (first entry) Matches 113; Conservative Barash SC, WPI; 2001-465566/50 N-PSDB; AAS41205. Local Similarity Sequence 128 AA

ij

```
Example 6; Fig 6; 646pp; English.
                                                                                                                                                                                                                           Escherichia coli polypeptide SEQ
                                                                                                                                        ABB52501 standard; Protein; 492
                                                                                                                                                                                                                                                                                                                                                                                                               12-MAR-2001; 2001WO-EP03445.
                                                                                                                                                                                                                                                                                                                                                                                                                                           2000FR-0003145.
2001FR-0001449.
                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bonacorsi S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2001-550253/61.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       492 AA;
                                                                                                                                                                                                                                                                                                                            Escherichia coli.
                                                                                                                                                                                                                                                                                                                                                       WO200166572-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                         10-MAR-2000;
02-FEB-2001;
                                                                                                                                                                                               11-FEB-2002
                                                                                                                                                                                                                                                                                                                                                                                   13-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            136;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Bingen E,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local S
Matches 136
                                                                                                                                                                    ABB52501;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                               15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        63
                                         495
                                                                   394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                120
                                                                                                                             ABB52501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Dp
                                          δ
                                                                   g
                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13;
                                                                                                                                                                                                                                                                                                                                                                                                                             13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                           New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell
                                                                                                                                                                                                                                                                                                                                                                                                                        capable of detecting 1000 or more genes from Drosophila. The invention useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABLi6176-ABL30511), expressed DNA sequences (ABL16176-ABL30511), expressed DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----IYTKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLTG--NPG-YSDWVPPQSFSNL 475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      140 LQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDTFFGS 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----- 494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The invention relates to an isolated nucleic acid detection reagent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LLGSGDYYTH-YKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQILASH---NP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sqplfllnhlaphaandddpmqapaeevsrfeyisnkthryyaamvsrldksvgsvlda
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       311 LKTYGFYNNSIIIYSSDNGGQPTAG----GSNWPLRGSKGTYWEGGIRAVGFVHSPLLK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       366 NKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDP-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TKPIFLYIAYQAVHS-----PLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLA
                          developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      94;
                                                                                                                                                                                                                                                                                                                                                                                 ID NO 16569; 21pp + Sequence Listing; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 486;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
Drosophila melanogaster polypeptide SEQ ID NO 16569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 556; DB 22;
Pred. No. 8.4e-44;
5; Mismatches 166;
                                                                                                                                                                                                                                                        EV.
                                                                                                                                                                                                                                                        Myers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GPNRWHNERITSSTGKSVW-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      65;
                                                                                                                                                                                                                                                        PWD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18.5%;
31.3%;
                                                                                                                                                    23-MAR-2001; 2001WO-US09231
                                                                                                                                                                                23-MAR-2000; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
                                                                                                                                                                                                                                                        ī
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                  Drosophila melanogaster
                                                                                                                                                                                                                                                       Adams M,
                                                                                                                                                                                                                                                                                2001-656860/75.
                                                                                                                                                                                                                           (PEKE ) PE CORP NY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  486 AA;
                                                                                                                                                                                                                                                                                               N-PSDB; ABL07362
                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; SEQ
                                         pharmaceutical.
                                                                                              WO200171042-A2
                                                                                                                                                                                                                                                                                                                                                       interactions
                           Drosophila;
                                                                                                                           27-SEP-2001
                                                                                                                                                                                                                                                       Venter JC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Sin
Matches 148;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            476
```

g

q ò

g

ò 셤 ò

ò

a

ò

NAME OF THE PROPERTY OF THE PR

```
15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The invention relates to a library of DNA fragments of Escherichia coll strains comprising polynucleotides (ABA88577-ABA88729 and ABA89533) and encoded proteins (ABB52459-ABB52914-ABB53094) of nature B2/D+A. The polynucleotides have potential antiinflammatory, antibacterial and immunosuppressive activity as part of pharmaceutical compositions used to treat, pallate or prevent extra-intestinal E collinfections. The polypeptides are useful for determining the phylogenic group of a given E. coli strain. These polypeptides can detect and treat an undesired development of E. coli, particularly an extra-intestinal infection that include systemic and non-diarrhoeal infections such as septicaemia, pyelonephritis and meningitis this is particularly advantageous as bacterial resistance is increasing with the more frequent use of broad spectrum antibiotics.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a library of DNA fragments of Escherichia coli strains for the phylogenic determination of a given strain comprises polynucleotides of nature 82/D+\ A-
| | : :| |||
336 dpraveyeelvrnt--svwlqlqqvsfgerniselrdqsriecpdpatgvkpclplegpc 393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Escherichia coli; B2/D+A-; antiinflammatory; antibacterial; immuosuppressive; extra intestinal infection; phylogeny; meningitis; systemic infection; non-diarrhoeal infection; septicaemia; pyelonephritis; antibiotic resistance.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GEKLEPS-TTSTSQPHLIFILADDQGFRDVGYHGSEI-KTPTLDKLAAEGVKLENYYV-Q 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHL 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                107;
                                                                                                                                                                                lfdleadpcersnlyaeyqnstifldlwsriqqfakqahppnnkpgdpncdpr 446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 492;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ပ်
                                                                                                            LENITADPYERVDLSNRYPG - - IVKKLLRRLSQFNKTAVPVRYPPKDPRSNPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        201; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Tinsley
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nassif X,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15.2%; Score 457.5; DB 26.2%; Pred. No. 2.1e-34.ive 75; Mismatches 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (INRM ) INSERM INST NAT SANTE & RECH MEDICALE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ID NO 350.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Clermont O,
                                                                                                                                                                                                                                                                                                                                                                                    Ā
```

g è	72	72 pisspsragiltgrmpfrtgir-swipsgkdvalgrneitianilkaggydtammgklhl 80GPVRKECMPTRRGEDTFFGSLIGGGDVVTHVKCNGBGMCGVDIVERDNAAANDONMG	130
. K	3		
Q	131	131 naggdrtdqpqaqdmgfdyslantagfvtdatldnakerprygmvyptgw	180
οy	237	237 YSTOMYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRS	284
q	181	lrngqptpradkmsgeyvssevvnwldnkkdskpfflyvaftevhsplaspkkyldmysq	240
οy	285	IININRRRYAAMLSCLDEAINNVTLALKTYGFYNNSIIIYS	325
đ	241	241 ymsayqkqhpdlfygdwadkpwrgvgeyyanisyldaqvgkvldkikamgeedntivift	300
οy	326	SDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCK	372
đ	301	sdngpvtrearkvyelnlagetdglrgrkdnlweggirvpaiikygkhlpggmvsd	356
ογ	373	EPVHITDMY PTLISLAEGIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPIYTKAKNG 432	432
QQ	357	357 tpvygldwmptlakmmnfklptdrtfdgeslvpvlegkalkrekplifgidmpfgddptd 416	416
δλ	433	SWAAG	492
q	417	417 ewidrnnkp 435	435
δy	493	493 VWLFNITADPYERVDLSNRYPGIVKKLLRRLSQFNKTAV 531	
qq	436	kjyniksdryetinligkkpdiekqmygkflky-ktdi 473	

Search completed: July 23, 2002, 19:02:50 Job time: 6936 sec

THIS PAGE BLANK (USPTO)

```
July 23, 2002, 19:04:14 ; Search time 22.71 Seconds (without alignments) 2327.131 Million cell updates/sec
                                                                                                                                                                                                                                                1 MAPRGCAGHPPPPSPQACVC.....VPVRYPPKDPRSNPRLNGGV
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                OM protein - protein search, using sw model
                                                                                                                                                                                                  US-09-495-823-7
550
                                                                                                                                                                                                        Title:
Perfect score:
                                                                                                                                                                                                                                                                                         Scoring table:
                                                                                                                                                                                                                                                    Sequence:
                                                                                                                        Run on:
```

283138 Total number of hits satisfying chosen parameters: 283138 seqs, 96089334 residues OLIGO Gapop 60.0 , Gapext 60.0 0 Word size Searched:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

summaries Post-processing: Listing first 45

PIR\_71:\* 1: pir1:\* 2: pir2:\* 3: pir3:\* Database :

pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES Result

Description	N-acetyldalactosam	probable thioredox	N-acetylgalactosam	N-acetylgalactosam	phosphoribosylamin	two-component resp	3-oxoacyl-[acyl-ca	hypothetical prote		hypothetical prote	Ψ	arylsulfatase (EC	_	hypothetical prote	>	SRB9 protein - yea	pr	•-	hypothetical prote	7	hypothetical prote	hypothetical prote		fimbrial protein s	septum-promoting G	hypothetical prote	probable glutathio	hypothetical prote	Q.
ID	154210	AI0374	KJHUAB	A44475	S76001	AC2385	A83230	AD1113	AH1474	T29618	807089	S01793	A37362	S61973	AH1535	B57062	AD2287	AG0271	E70327	H72468	H84014	AI2450	AH1524	C39415	T45541	T49590	C83292	F84685	AH0252
DB	2	7	-	-	7	~	7	7	7	7	7	~	~	~	7	~	7	~	7	~	7	7	~	7	~	7	~	7	7
Length	473	289	533	535	176	219	330	381	381	452	551	551	567	647	1216	1420	82	108	114	123	123	155	159	180	198	200	220	229	232
Query Match	2.0	1.8	•	•	1.5	1.5	1.5	1.5	•	•	1.5	•	1.5	1.5	1.5	1.5	1.3	1.3	1.3	•	1.3	•	1.3	•	•	٠	1.3	1.3	1.3
Score	11	10	10	10	8	άο	80	80	80	ထ	80	80	8	80	80	80	7	7	7	7	7	7	7	7	7	7	7	7	7
esult No.	1	7	m	4	S	9	7	œ	6	10	=	12	13	14	15	16	17	18	19	70	21	22	23	24	25	56	27	78	53

precorrin-2 C20-me US4 protein - huma hypothetical prote		probable aldolase hypothetical prote probable carbon-ni	33kD chaperonin, h hypothetical prote hypothetical prote	tRNA pseudouridine glycosyl transfera
AC0758 QQBE74 T20432	A12175 AD3548 G81148 E91262	H86102 A12390 AB0115	AF2004 S75112 T15835	B83951 A98140
212	0000	0000	9000	00
237 238 241	248 277 278 286	2222 2222 2226 246	301 302 303	304 319
11.1	 			1.3
7 7	~~~			7
30 31 32	33 35 35	33 38 39	4 4 4 4 3 4 2 1 5	44 45 :

## ALIGNMENTS

```
Nacetylgalactosamine-4-sulfatase (EC 3.1.6.12) precursor - rat (fragment)
N.Alternate names: arylsulfatase (EC 3.1.6.1) B [misidentification]
C.Species: Rattus norvegicus (Norway rat)
C.Sate: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 02-Jun-2000
C.Accession: 154210
R.Kunieda, T.
Genomics 29, 582-587, 1995
A.Title: Mucopolysaccharidosis type VI in rats: Isolation of cDNAs encoding arylsulfa A; Reference number: 154210; MUD:96121368
A.Tette: Mucopolysaccharidosis type VI in rats: Isolation of cDNAs encoding arylsulfa A; Reference number: 154210
A; Reference number: 154210
A; Restques: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: mRNA
A; Residues: 1-473 <RES>
A; Cross-references: GB:D49434; NID:91065603; PIDN:BAA08412.1; PID:91089794
C; Genetics:
A; Gene ARSB
C; Superfamily: animal sulfatase
C; Keywords: sulfuric ester hydrolase
F; 31/Modified site: 3-oxoalanine (Cys) #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 473;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity 100.0%; Score 11; DB 2; Similarity 100.0%; Pred. No. 0.012; 1; Conservative 0; Mismatches 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Simi
Matches 11;
RESULT
                                  154210
```

STHMVGKWHLG 180 STHMVGKWHLG 89 170 ŏ

4

Dp

RESULT

Probable thioredoxin YPO3082 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001
C;Accession: A10374
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M. deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G. il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, R.; Whitehead, S.; Barrel A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: A10374
A;Atatus: preliminary
A;Molecule type: DNA
A;Residues: 1-289 <KUR>
A;Residues: 1-289 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC92324.1; PID:g15981035; GSPDB:GN00175
C;Genetics:
A;Gene: YPO3082

24 10:05:45 2002

Wed Jul

```
10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local
Matches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ó
                                                                                                                                                                                                                                                                                                                                                                                                                         Nacetylgalactosamine-4-sulfatase (EC 3.1.6.12) precursor [validated] - human Naterinate names: arylaulfatase B (ASB); chondroitinase; chondroitinsulfatase; G4S; N-6 C; Species: Homo sapiens (man) C; Date: 31-Dec-1993 #sequence_revision 27-oct-1995 #text_change 08-Dec-2000 C; Accession: S35990; S33307; A35078; A45659; A42449; B42449; C42449; I54217; A56865 R; Peters. C.W.B. C.W.B. Submitted to the EMBL Data Library, March 1993 A; Reference number: S35990 A; Molecule type: DNA A; Reference s35990 A; Molecule type: DNA A; Residues: 1-533 < PET. Submitted to the EMBL: X72735; NID:9289009; PIDN:CAA51272.1; PID:9825628 R; Modaressi, S.; Rupp, K.; von Figura, K.; Peters, C. Biol. Chem. Hoppe-Seyler 374, 327-335, 1993 A; Title: Structure of the human arylsulfatase B gene. A; Reference number: S33307; MUD:93332648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type: DNA
A; Residues: 1-104 <MOD>
A; Residues: 1-104 <MOD>
A; Residues: 1-104 <MOD>
A; Cross-references: EMBL:X72735; EMBL:X72736; EMBL:X72737; EMBL:X72738; EMBL:X72739; EME
A; Cross-references: EMBL:X72735; EMBL:X72739; EMBL:X72739; EME
A; Note: the enzyme is referred to as EC 3.1.6.9
R; Peters, C.; Schmidt, B.; Rommerskirch, W.; Rupp, K.; Zuehlsdorf, M.; Vingron, M.; Meye
A; D Blol. Chem. 265, 3374-3381, 1990
A; Title: Phylogenetic conservation of arylsulfatases, cDNA cloning and expression of hum
A; Reference number: A35078; MUID:90153994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Cross-references: GB:S57777; NID:9236697; PIDN:AAB19988.1; PID:9236698
A; Note: sequence extracted from NCBI backbone (NCBIN:57777, NCBIP:57778)
A; Note: the enzyme is referred to as EC 3.1.6.1
A; Note: parts of this sequence, including the amino end of the mature protein, were deterned form is described with a proteolytic cleavage somewhere between residue 450 and R; Jin, Wu.D; Jackson, C.E.; Desnick, R.J.; Schuchman, E.H.
Am. J. Hum. Genet. 50, 795-800, 1992
A; Ittle: Mucopolysaccharidosis type VI: identification of three mutations in the arylsul
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          isolation of
                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Wolecule type: mRNA
A; Residues: 1-357, 'V', 359-533 < PE2>
A; Cross-references: 6B:J05225, NID:9179076; PIDN:AAA51784.1; PID:9179077
A; Note: parts of this sequence were determined by protein sequencing
A; Note: the enzyme is referred to as EC 3.1.6.1
B; Litjens, T.; Morris, C.P.; Gibson, G.J.; Beckmann, K.R.; Hopwood, J.J.
Biochem. Int. 24, 209-215, 1991
A;Title: Human N-acetylgalactosamine-4-sulphatase: protein maturation and A; Reference number: A45559; MUID:92028992
A; Accession: A45559
A; Molecule type: DNA; protein
A; Residues: 1-104 < LIT>
                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: mRNA
A; Residues: 115-116, 'R', 118 <JIN>
A; Cross-references: GB:S90729; NID:g247486; PIDN:AAB21831.1; PID:g247487
A; Note: sequence extracted from NCBI backbone (NCBIN:90729, NCBIP:90731)
A; Accession: B42449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Molecule type: mRNA
A; Residues: 234-235, 'P', 237-238 <JI2>
A; Cross-references: GB:S90736; NID:g247488; PIDN:AAB21832.1; PID:g247489
A; Note: sequence extracted from NCBI backbone (NCBIN:90736, NCBIP:90739)
A; Accession: C42449
                                                                                                                ö
                                         Length 289;
                                                                                                             Indels
Score 10; DB 2; Ler
Pred. No. 0.085;
           1.8%; Scc. 100.0%; Pred. No. v. 0. Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Reference number: A42449; MUID:92197625
A; Accession: A42449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Molecule type: mRNA
A; Residues: 403-404, 'Y', 406-407 <JI3>
                                                                   Best Local Similarity 100.
Matches 10; Conservative
                                                                                                                                                                                100 TPTLDKLAAE 109
                                                                                                                                                                                                                             42 TPTLDKLAAE
                                         Query Match
                                                                                                                                                                                                                                                g
                                                                                                                                                                                    ò
```

```
A,Cross-references: GB:590743; NID:9247490; PIDN:AAB21813.1; PID:9247491
A,Notes: sequence extracted from Nerb Backbook (NEBN:90743, NCBP:90747)
A,Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to as CC 3.16.1
A;Note: the enzyme is referred to a control 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-acetylgalactosamine-4-sulfatase (EC 3.1.6.12) precursor - cat
N-acetylgalactosamine-4-sulfatase (EC 3.1.6.12) precursor - cat
N-acetylgalactosamine-4-sulfatase (ARSB); chondroitinase; chondroitinsulfatase; G4S;
C; Species: Felis silvestris catus (domestic cat)
C; Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C; Accession: A4475
R; Jackson, C.E.; Yuhki, N.; Desnick, R.J.; Haskins, M.E.; O'Brien, S.J.; Schuchman, E
Genomics 14, 403-411, 1992
A; Title: Feline arylsulfatase B (ARSB): isolation and expression of the cDNA, compari
A; Reference number: A44475
A; Accession: A44475
A; Status: preliminary; not compared with conceptual translation
A; Molecule type: mRNA
A; Residues: 1-535 <JAC>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1.8%; Score 10; DB 1; Length 533; 100.0%; Pred. No. 0.15; 0; Indels ive 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             140 THMVGKWHLG 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  171 THMVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
```

Wed Jul 24 10:05:45 2002

```
Applies

Appropriation of the protein lime of the protein of 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa A;Reference number: A82950; MUID:20437337
A;Accession: A83230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Cross-references: GB:AE004755; GB:AE004091; NID:g9949456; PIDN:AAG06721.1; GSPDB:GN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3-oxoacy1-[acy1-carrier-protein] synthase (EC 2.3.1.41) III PA3333 [similarity]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P.; Hickey,
A.; Larbig,
                                                                                                     A;Molecule type: DNA
A;Residues: 1-219 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB76334.1; PID:917133772; GSPDB:GN00179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 02-Feb-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C;Accession: A83230
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, .; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                           A:Experimental source: strain PCC 7120
C;Genetics: all4635
C;Superfamily: regulatory protein comA; response regulator homology
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 330;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 219;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C;Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase III
C;Reywords: acyltransferase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1.5%; Score 8; DB 2;
100.0%; Pred. No. 11;
tive 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1.5%; Score 8; DB 2;
100.0%; Pred. No. 7.3;
iive 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Experimental source: strain PAO1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 100.0
Local 8; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 1.5
Best Local Similarity 100.
Matches 8; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Molecule type: DNA
A; Residues: 1-330 <STO>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 513 PGIVKKLL 520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           127 PGIVKKLL 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Gene: fabH2; PA3333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Status: preliminary
                                                          A; Status: preliminary
     A; Accession: AC2385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              53 EEEGALLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C; Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002 C; Accession: ACJ385 R; Kaneko, T; Wakamura, Y; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S DNA Res. 8, 205-213, 2001 A; Tabata, S A; Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Analy, Reference number: AB1807; MUID:21595285; PMID:11759840
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A.Accession: S7601
A.Status: nucleic acid sequence not shown; translation not shown
A.Status: nucleic acid sequence not shown; translation not shown
A.Status: nucleic acid sequence not shown; translation not shown
A.Molecule type: DNA
A.Coss-references: EMBL:D64006; GB:AB001339; NID:g1001291; PIDN:BAA10848.1; PID:g100136
A.Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C.Sepertics:
A.Gene: purE
C.Superfamily: phosphoribosylaminoimidazole carboxylase catalytic chain; phosphoribosyla
C.Superfamily: phosphoribosylaminoimidazole carboxylase catalytic chain, nonology <PCC>
F.6-1397Domain: phosphoribosylaminoimidazole carboxylase catalytic chain homology <PCC>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  phosphoribosylaminoimidazole carboxylase (EC 4.1.1.21) catalytic chain - Synechocystis s N;Alternate names: protein sl10901 (Speciaes: Synechocystis sp. A;Variety: PCC 6803 (Speciaes: Synechocystis sp. A;Variety: PCC 6803 (Speciaes: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000 (Spate: 25-Apr-1997 #sequence_revision 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000 (Spate: 25-Apr-1997 #sequence_revision 25-Apr-1997 #sequence_
                                                                          A:Note: sequence extracted from NCBI backbone (NCBIP:117976)
C:Comment: This enzyme is frequently misidentified as EC 3.1.6.1.
C:Punction:
A:Pescription:
A:Pescription: hydrolyzes N-acetylgalactosamine-4-sulfate units in chondroitin sulfate a C;Superfamily: animal sulfatase
F:14-42/Domain: signal sequence #status predicted <SIG>F:43-42/Domain: signal sequence #status predicted <AMAT>
F:46-46/Product: alpha chain #status predicted <AMAT>
F:468-535/Product: beta chain #status predicted <AMAT>
F:93/Modified site: 3-oxoalanine (Cys) #status predicted
F:190,281,428,460/Binding site: carbohydrate (Asn) (covalent) #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Superfamily: phosphoribosylaminoimidazole carboxylase catalytic chain; phosphoribosyla; Keywords: carbon-carbon lyase; carboxy-lyase; purine nucleotide biosynthesis; 6-139/Domain: phosphoribosylaminoimidazole carboxylase catalytic chain homology <PCC>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               two-component response regulator all4635 [imported] - Anabaena sp. (strain PCC 7120)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
A; Cross-references: GB:S48472; NID:q258855; PIDN:AAB23941.1; PID:q258856
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 1; Length 535; 0.15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 176;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1.5%; Score 8; DB 2;
100.0%; Pred. No. 6;
ative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 1.8%; Score 10; DB Best Local Similarity 100.0%; Pred. No. 0.1 Matches 10; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Reference number: S74322; MUID: 97061201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 171 THMVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            142 THMVGKWHLG 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C;Species: Anabaena sp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    248 QILASHNP 255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |||||||||
131 QILASHNP 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

M.J.; K.; L

ö

- Pse

C; Genetics

σ

ò g

```
arylsulfatase (EC 3.1.6.1) precursor - sea urchin (Hemicentrotus pulcherrimus)
C;Species: Hemicentrotus pulcherrimus
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 26-Aug-1999
C;Accession: S01793 #sequence_revision 30-Sep-1991 #text_change 26-Aug-1999
C;Accession: S01793 #sequence and expression of the gene for arylsulfata A;Reference number: S01793; MUID:89030699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A37362 arylsulfatase (EC 3.1.6.1) precursor - sea urchin (Strongylocentrotus purpuratus) (Species: Strongylocentrotus purpuratus (purple urchin) C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Nov-1999 C;Accession: A37362
                                                                                                                            Cypecies: Lytechinus pictus (painted urchin) (Lytechinus pictus)
Cypecies: Lytechinus pictus (painted urchin)
Cypecies: Lytechinus pictus
Cypecession: S07089
R:Yamada, K.; Akasaka, K.; Shimada, H.
Eur. J. Biochem. 186, 405-410, 1989
A;Title: Structure of sea-urchin arylsulfatase gene.
A;Reference number: S07089; MUID:90092130
A;Reference number: S07089
A;References: EMBL:X16679
C;Genetics:
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A Molecule type: mRNA
A; Residues: 1-551 < SAS>
A; Cross-references: EMBL:X17015; NID:99432; PID:99433
A; Note: part of this sequence was confirmed by protein sequencing A; Note: the authors translated the codon CAC for residue 61 as Asp C; Superfamily: animal sulfatase
C; Superfamily: animal sulfatase
C; Keywords: sulfuric ester hydrolase
C; Keywords: sulfuric ester hydrolase
E; 1-20/Domain: signal sequence #status predicted < SIG>
E; 21-551/Product: arylsulfatase #status predicted < MATP>
F; 100/Modified site: 3-oxoalanine (Cys) #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 551;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 551;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Introns: 103/2; 160/1; 197/3; 272/1; 327/1
C; Superfamily: animal sulfatase
C; Keywords: sulfuric ester hydrolase
F;100/Modified site: 3-oxoalanine (Cys) #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1.5%; Score 8; DB 2;
100.0%; Pred. No. 17;
ative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1.5%; Score 8; DB 2
100.0%; Pred. No. 17;
Live 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 1.5
Best Local Similarity 100.
Matches 8; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Matches 8; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               173 MVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              134 MVGKWHLG 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             173 MVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               153 MVGKWHLG 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Accession: S01793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Gene: Ars
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            13
                                                                                11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Righaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker J. Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H. D.; Jones, L.M.; Kats, U. Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma A;Authors: Kreft, J.; Sinoes, N.; Taterrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-381 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC95568.1; PID:g16412764; GSPDB:GN00178
A;Experimental source: strain Clipil262
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Accession: T29618
R;Du, Z.; Leimbac, D.
Submitted to the EMBL Data Library, March 1996
S;Du, Z.; Leimbac, D.
Submitted to the EMBL Data Library, March 1996
S;Description: The sequence of C. elegans cosmid D1014.
A;Reference number: 220652
A;Reference number: 220653
A;Status: prelliminary; translated from GB/EMBL/DDBJ
A;Rolecule type: DNA
A;Residues: 1-452 <DUZ>
A;Cross-references: EMBL:U53180; PIDN:AAA96290.1; GSPDB:GN00023; CESP:D1014.1
A;Experimental source: strain Bristol N2; clone D1014
                                                                                                                                                                                                                                                                                                                                                                              hypothetical protein lin0335 [imported] - Listeria innocua (strain Clip11262)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         hypothetical protein D1014.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
                                                                                                                                                                                                                                                                                                                                                                                                       C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ;
0
                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 381;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 452;
                                                                             Length 381;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0; Indels
                                                                                                                                  ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1.5%; Score 8; DB 2;
100.0%; Pred. No. 12;
tive 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1.5%; Score 8; DB 2;
100.0%; Pred. No. 14;
tive 0; Mismatches
                                                                                DB 2;
                                                                                              100.08; Pred. w.
                                                                             1.5%; Score 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Introns: 256/3; 336/2; 376/3; 405/3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 100.00
Loc 8; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 100.(
Matches 8; Conservative
                                                                                                                                  Conservative
                                                                                                         Best Local Similarity
Matches 8; Conserv
                                                                                                                                                                                                                     181 GFYNNSII 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          315 GFYNNSII 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 GFYNNSII 188
                                                                                                                                                                                       315 GFYNNSII 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               173 MVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Gene: CESP:D1014.1
A;Map position: 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                         C; Accession: AH1474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Accession: AH1474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Gene: 11n0335
A; Gene: 1mo0307
                                                                                Query Match
```

RESULT

ò

ð g

ö

```
δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AH1339
Pyrvate-flavodoxin oxidoreductase homolog nifJ [imported] - Listeria innocua (Strain C] C; Species: Listeria innocua (Species: 127-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001 (Species: NH1535) (Species: Species: Species: NH1535) (Species: Species: Specie
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Whothetical protein YPL093w - yeast (Saccharomyces cerevisiae)
Nypothetical protein LPG15w
Nypothetical protein LPG15w
C; Species: Saccharomyces cerevisiae
C; Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 05-Nov-1999
C; Accession: S61973
R; Wang, Y; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Hall, J.; Storms, R.K.; Vc submitted to the EMBL Data Library, December 1995
S; Date: 10-Apr-1996 #sequence of Saccharomyces cerevisiae chromosome XVI left arm.
A; Reference number: S61973
A; Reference number: S61973
A; Residues: 1-647 < VANA>
A; Residues: 1-647 < VANA>
A; Residues: BMBL:U43281; NID:g1151218; PID:g1151233; GSPDB:GN00016; MIPS:YPL093x
A; Map position: 16L
R; Yang, Q.; Angerer, L.M.; Angerer, R.C.
Dev. Biol. 1135, 53-65, 1989
A;Title: Structure and tissue-specific developmental expression of a sea urchin arylsulf A;Reference number: A37362; MUID:89357267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                       A;Cross-references: GB:M28404; GB:M25815; NID:g161440; PID:g161441
C;Superfamily: animal sulfatase
C;Keywords: sulfuric ester hydrolase
F;115/Modified site: 3-oxoalanine (Cys) #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 647;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 567;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1.5%; Score 8; DB 2;
100.0%; Pred. No. 17;
Live 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1.5%; Score 8; DB 2;
100.0%; Pred. No. 19;
iive 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
Matches 8; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
Matches 8; Conserv
                                                                                                                                                                                                                    A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-567 <YAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               173 MVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            168 MVGKWHLG 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 48 QALEEEEE 55
                                                                                                                                                                            A; Accession: A37362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
```

```
A; Gene: nifJ
C; Superfamily: pyruvate (flavodoxin) dehydrogenase; ferredoxin 2[4Fe-4S] homology
Query Match
Dest Local Similarity
Ouery Matches 8; Score 8; DB 2; Length 1216;
Best Local Similarity
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 405 ErisEGLR 412
Db 902 ErisEGLR 909
Search completed: July 23, 2002, 19:09:17
Job time: 303 sec
```

. د د sapien

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a alicense agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                           chlamydia t
                                                                                     homo sapien
thermococcu
               coturnix co
                                            rattus norv
                                                          pisum sativ
                                                                        schizosacch
                                                                                                                    drosophila
                                                                                                                                                 mesostigma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "Mucopolysaccharidosis type VI in rats: isolation of cDNAs encoding arylsulfatase B, chromosomal localization of the gene, and identification of the mutation."; denomics 29:582-587(1995).

-!- CATALYTIC ACTIVITY: Hydrolysis of the 4-sulfate groups of the Nacetyl-D-galactosamine 4-sulfate units of chondroitin sulfate and dermatan sulfate.

-!- SUBUNIT: HOMODIMER (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Chórdata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SUBCELLULAR LOCATION: Lysosomal.
DISEASE: DEFECTS IN AREB ARE THE CAUSE OF MAROTEAUX-LAMY
SYNDROME; ALSO KNOWN AS MUCOPOLYSACCHARIDOSIS TYPE VI (MPS-VI).
SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
                                                                                                                                  homo
                                                                                                                                                               уошо
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                 P29843
Q9ulx6
Q9muk8
             Q02916
084290
P15589
082043
Q9uug8
Q9h2c0
Q56303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2-AMINO-3-OXOPROPIONIC ACID (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                           01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
Arylsulfatase B (EC 3.1.6.12) (ASB) (N-acetylgalactosamine-4-sulfatase) (G4S) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; FALSE_NEG.
Hydrolase; Glycoprotein; Lysosome; Mucopolysaccharidosis.
                                                                                                                                                                                                                                                                                  473 AA
                                                                                                                                                                                                        ALIGNMENTS
             NFL_COTJA
Y288_CHLTR
STS_RAT
ILV5_PEA
TU12_SCHPO
GAN_HUMAN
                                                                                                                  HS7A_DROME
NA95_HUMAN
NU5C_MESVI
                                                                                                                                                              ILF1_HUMAN
                                                                                                      FOR_THELI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-SPRAGUE-DAWLEY; TISSUE-Liver;
MEDLINE-96121368; Pubmed-8575749;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR000917; Sulfatase.
Pfam; PF00884; Sulfatase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D49434; BAA08412.1; -. HSSP; P15848; 1FSU.
                                                                                                                                                                                                                                                                                  STANDARD;
5551
5553
557
577
581
621
641
645
652
                                                                                                                                                                                                                                                                                                                                                                                                      Rattus norvegicus (Rat)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      87
461
95
132
387
128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      87
57
61
121
345
128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hydrolase; (NON_TER MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Kunieda T.;
r
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DISULFID
DISULFID
CARBOHYD
                                                                                                                                                                                                                                                                               ARSB_RAT
P50430;
                                                                                                                                                                                                                                                                                                                                                                                       ARSB.
48888844444
48888844444
5487888
                                                                                                                                                                                                                                                                                  (without alignments)
1303.290 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              neisseria m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       homo sapien
caulobacter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gall
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         human cytom
human cytom
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               P50430 rattus norv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         strongyloce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  herpes simp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          bacillus ha
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   streptococc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          arabidopsis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             felis silve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hemicentrot
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       saccharomyc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     saccharomyc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                acanthamoeb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              synechocyst
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    mycobacteri
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gallus gall
klebsiella
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        norv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             synechocyst
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    epstein-bar
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   salmonella
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     rattus
gallus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Description
                                                                                                                                                                                        MAPRGCAGHPPPPSPQACVC......VPVRYPPKDPRSNPRLNGGV
                                                                                                 July 23, 2002, 19:07:55; Search time 16.34 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      P14 600
P50473
P50473
P50473
Q05593
P05684
P90513
P7310
Q14320
Q14320
Q14320
Q14320
Q14320
Q1673
P07310
Q1673
P07310
Q1673
P07310
P0731
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q42569
P36837
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 P34059
P03210
Q61211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     P23565
                           Compugen Ltd
                                                                                                                                                                                                                                                                                                                       Total number of hits satisfying chosen parameters:
             GenCore version 4.5
Copyright (c) 1993 - 2000 Comp
                                                                                                                                                                                                                                                               105224 seqs, 38719550 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MML6_MYCTU
COMB_STRPN
MUC_HUMAN
C6ST_CHICK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EGRZ_MOUSE
C901_ARATH
YHIP_ECOLI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ARSB_HUMAN
ARSB_FELCA
PUR6_SYNY3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AINX_RAT
CPT7_CHICK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NOG1_YEAST
SRB9_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        XAP5_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GA6S_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                   summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PANC_NEIMB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HSLO_SYNY3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRUB_BACHD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CAUCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MUCB_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VIE1_HCMVA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       _HCMVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LIGA_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ARS_HEMPU
ARS_STRPU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PSA3_ACACA

    protein search, using sw model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /GLG_HSV11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ARS_KLEAE
                                                                                                                                                                                                                    OLIGO
Gapop 60.0 , Gapext 60.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RECF_
                                                                                                                                                                                                                                                                                                                                                                                                   Post-processing: Listing first 45
                                                                                                                                                                                                                                                                                                                                                      Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                           US-09-495-823-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB
                                                                                                                                                                                                                                                                                                                                                                                                                                 SwissProt_40:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    %
Query
Match Length D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score
                                                                                                                                                           Title:
Perfect score:
                                                                                                                                                                                                                       Scoring table:
                                                                                                                                                                                                                                                                                             Word size :
                                                                     OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                 Database :
                                                                                                                                                                                        Sequence:
                                                                                                                                                                                                                                                                 Searched:
                                                                                                   Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Š
```

```
RESULT 3
ARSB_HUMAN
ID ARSB_HUMAN
    CARBOHYD
CARBOHYD
CARBOHYD
                                                                                       NON_TER
SEQUENCE
                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                 P15848;
                                                                                                                                                                                                                                                                                                  38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         and
        FT
FT
FT
SO
                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE OF 34-285 FROM N.A.

MEDLINE-96332200; PubMed-8710849;

MEDLINE-96332200; PubMed-8710849;

Byers M., Safftig P., Schmidt P., Hafner A., McLoghlin D.B.,

Schmahl W., Hess B., von Figura K., Peters C.W.B.;

"Targeted disruption of the arylsulfatase B gene results in mice resembling the phenotype of mucopolysaccharidosis VI.";

Proc. Natl. Acad. Sci. U.S.A. 93:8214-8219(1996).

-! CAPALYIT ACTIVITY: Hydrolysis of the 4-sulfate groups of the N-acetyl-D-galactosamine 4-sulfate units of chondroitin sulfate and dermatan sulfate.

-! SUBUNIT: HOMODIMER (BY SIMILARITY).

-! SUBUNIT: HOMODIMER (BY SIMILARITY).

-! SUBCELLUAR LOCATION: Lysosomal.

-! SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 1-33 FROM N.A.
MEDILTE-222486; Pubmed-1572648;
Grompe M., Pleretti M., Caskey C.T., Ballabio A.;
"The sulfatase gene family: cross-species PCR cloning using the MOPAC
                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                             . .) (POTENTIAL). . . . (POTENTIAL). . . . (POTENTIAL).
                          (POTENTIAL)
                                                                                                                                                                                                                 ó
                                                                                                                                                                                                                                                                                                                                                                                                                                              01-OCT-1996 (Rel. 34, Created)
1-OCT-1996 (Rel. 34, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Arylsulfatase B (EC 3.1.6.12) (ASB) (N-acetylgalactosamine-
                                                                                                                                                                        Length 473;
                                                                                                                                                                   2.0%; Score 11; DB 1; Length 473
100.0%; Pred. No. 0.012;
iive 0; Mismatches 0; Indels
                                                                                                           4E114C923A24AF8F CRC64;
                                                               (GLCNAC.
                                                (GLCNAC.
                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                         285 AA
                     N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROSITE; PS00523; SULFATASE_1; PARTIAL.
PROSITE; PS00149; SULFATASE_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Glycoprotein; Lysosome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR000917; Sulfatase.
Pfam; PF00884; Sulfatase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1-sulfatase) (G4S) (Fragments).
                                                                                                             53320 MW:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; M82877; AAA37261.1; -. EMBL; X92096; CAA63067.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Genomics 12:755-760(1992).
                                                                                                                                                                                           Best_Local Similarity 100.
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     34
45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MGD; MGI:88075; As1-s.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mus musculus (Mouse).
                                                                                                                                                                                                                                                        170 STHMVGKWHLG 180
219
231
355
366
398
473 AA;
                                                                                                                                                                                                                                                                                                  79 STHMVGKWHLG 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    P15848; 1FSU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ARSB OR AS1
                                                                                                                                                                                                                                                                                                                                                                                                         ARSB_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hydrolase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    technique.
                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NON_CONS
ACT_SITE
                        CARBOHYD
                                             CARBOHYD
    CARBOHYD
                                                                                                           SECUENCE
                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NON_TER
                                                                                                                                                                                                                                                                                                                                                               RESULT 2
ARSB_MOUSE
  FT
FT
FT
SO
                                                                                                                                                                                                                                                           ô
                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                 PAGE STANDAR SECOND COLOR SERVING SERVING SECOND SE
```

```
MEDLINE-90152677; PubMed-1968043; Schuchman B.H., Jackson C.E., Desnick R.J.; Schuchman B.H., Jackson C.E., Desnick R.J.; Human arylsulfatase B: MOPAC Cloning, nucleotide sequence of a full-length cDNA, and regions of amino acid identity with arylsulfatases A length cDNA, and regions of amino acid identity with arylsulfatases A
                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (N-acetylgalactosamine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-95354208; PubMed-7628016; Schmidt B., Selmer T., Ingendoh A., von Figura K.; An ovel amino acid modification in sulfatases that is defective in multiple sulfatase deficiency."; Cell 82:271-278(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     [4] SEQUENCE OF 1-104 FROM N.A. SEQUENCE OF 1-104 FROM N.A. MEDLINE-92028992; PubMed-1930244; Litjens T., Morris C.P., Gibson G.J., Beckmann K.R., Hopwood J.J.; Litjens T., Morris C.P., Gibson G.J., Beckmann N-acetylgalactosamine-4-sulphatase: protein maturation and "farman N-acetylgalactosamine-1-sulphatase:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo saplens (Human).
Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostoml;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
(GLCNAC. ..) (POTENTIAL). (GLCNAC. ..) (POTENTIAL). (GLCNAC. ..) (POTENTIAL). (GLCNAC. ..) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-97184692; PubMed-9032078;
Bond C.S., Clements P.R., Ashby S.J., Collyer C.A., Harrop S.J.,
Hopwood J.J., Guss J.M.;
"Structure of a human lysosomal sulfatase.";
                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
MEDLINE-90153994; PubMed-2303452;
Peters C., Schmidt B., Rommerskirch W., Rupp K., Zuehlsdorf M., Vingron M., Meyer H.E., Pohlmann R., von Figura K.;
"Phylogenetic conservation of arylsulfatases. cDNA cloning and expression of human arylsulfatases B.";
Blol. Chem. 265:3374-3381(1990).
                                                                                                                                                                                                                                    1.8%; Score 10; DB 1; Length 285; 100.0%; Pred. No. 0.073;
                                                                                                                                                                                                                                                                                              0; Indels
                                                                                                                                                  36BA5B98EA0770C9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Modaressi S., Rupp K., von Figura K., Peters C., "Structure of the human arylsulfatase B gene."; Biol. Chem. Hoppe-Seyler 374:327-335(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-APR-1990 (Rel. 14, Created)
01-APR-1990 (Rel. 14, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
Arysulfattase B precursor (EC 3.1.6.12) (ASB)
4-sulfatase) (G4S).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2-AMINO-3-OXOPROPIONIC ACID MODIFICATION SITE
                                                                                                                                                                                                                                                                                           0; Mismatches
N-LINKED (
N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A. MEDLINE=93332648; PubMed=7687847;
                                                                                                                                                  ММ,
                                                                                                                                                  31727
                                                                                                                                                                                                                                    Query Match 1.8
Best Local Similarity 100.
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genomics 6:149-158(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
86
177
189
264
285
                                                                                                                                                                                                                                                                                                                                                        171 THMVGKWHLG 180
86
177
189
264
285
                                                                                                                                                                                                                                                                                                                                                                                          THMVGKWHLG 47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
```

us-09-495-823-7.oli.rsp

JOINED

```
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VARIANT MPS-IV ARG-302.

Villani G.R.D., Balzano N., di Natale P.;

"Two novel mutations of the arylsulfatase B gene in two Italian patients with severe form of mucopolysaccharidosis.";

Hum. Mutat. 11:410-410(1998).

-!- CATALYTIC ACTIVITY: Hydrolysis of the 4-sulfate groups of the Nacetyl-D-galactosamine 4-sulfate units of chondroitin sulfate and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SYNDROME; ALSO KNOWN AS MUCOPOLYSACCHARDOSIS TYPE UT (MPS-VI).

THIS DISEASE IS CHARACTERIZED BY THE ACCUMULATION OF DERWATAN
SULFATE IN LYSOSOMES. CLINICAL FEATURES CAN INCLUDE ABROATAN
GROWTH, SHORT STATURE, STIFF JOINTS, SKELETAL MALFORMATIONS,
CORNEAL CLOUDING, HEPATOSPLENOMEGALY, AND CARDIAC ABNORMALITIES.

M WIDE WARRATION IN CLINICAL SEVERITY IS OBSERVED.

DISEASE: MULTIPLE SULFATASE DEFICIENCY (MSD) IS A DISORDER THAT
COMBINES FEATURES OF METACHROMATIC LEUKODYSTROPHY AND OF
MUCOPOLYSACCHARIDOSIS; IT IS CHARACTERIZED BY A DECREASED
ACTIVITY OF ALL KNOWN SULFATASES. IT SEEMS TO BE CAUSED FROM THE
LACK OF POST-TRANSLATIONAL MODIFICATION OF A CYSTEINE INTO 2-
AMINO-3-OXOPROPIONIC ACID.
                                                              Jin W.-D., Jackson C.E., Desnick R.J., Schuchman E.H.; "Mucopolysaccharidosis type VI: identification of three mutations in the arylsulfatase B gene of patients with the severe and mild phenotypes provides molecular evidence for genetic heterogeneity."; Am. J. Hum. Genet. 50:795-800(1992).
                                                                                                                                                                                                                            "Mucopolysaccharidosis VI (Maroteaux-Lamy syndrome). An intermediate calinical phenotype caused by substitution of valine for glycine at position 137 of arylsulfatase B.".
J. Biol. Chem. 266:21386-21391(1991).
                                                                                                                                                                                                                                                                                                                        MEDLINE-96213747; PubMed-8651289;
Litjens T., Brooks D.A., Peters C., Gibson G.J., Hopwood J.J.;
Identification, expression, and biochemical characterization of N-
acetylgalactosamine-4-sulfatase mutations and relationship with
clinical phenotype in MPS-VI patients.";
Am. J. Hum. Genet. 58:1127-1134(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          novel mutant alleles of the arylsulfatase B gene in two
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Voskoboeva E., Isbrandt D., von Figura K., Krasnopolskaya X.,
                                                                                                                                                                                                                                                                                                            VARIANTS MPS-VI MET-92; GLN-95; CYS-210; PRO-393 AND PRO-498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUBCELLULAR LOCATION: Lysosomal.
DISEASE: DEFECTS IN ARSB ARE THE CAUSE OF MAROTEAUX-LAMY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           patients with intermediate form of mucopolysaccharidosis VI
(Maroteaux-Lamy syndrome).";
                                                                                                                                                                              MEDLINE-92042029; PubMed-1718978;
Wicker G., Prill V., Brooks D., Gibson G., Hopwood J.,
von Figura K., Peters C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
                              VARIANTS MPS-VI ARG-117; PRO-236 AND TYR-405.
MEDLINE-92197625; PubMed=1550123;
                                                                                                                                                               VARIANT MPS-VI VAL-137, AND VARIANT MET-376.
                                                                                                                                                                                                                                                                                                                                                                                                                                          VARIANTS MPS-VI TRP-152 AND GLN-160.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-94171224; PubMed-8125475;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; J05225; AAA51784.1; -.
EMBL; M32373; AAA5179.1; -.
EMBL; X72735; CAA51272.1; -.
EMBL; X72736; CAA51272.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (Maroteaux-Lamy syndrome).";
Hum. Genet. 93:259-264(1994).
Structure 5:277-289(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      dermatan sulfate.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Peters C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Four
```

```
ö
                                                                                                                                                                                                                                                                                                                                                  N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .).
T-LINKED (GLCNAC. .).
N-LINKED (GLCNAC. .).
T-SM (IN MPS-VI; MILD FORM).
T-> Q (IN MPS-VI; MILD/SEVERE FORM).
T-> Q (IN MPS-VI; MILD/SEVERE FORM).
T-> Q (IN MPS-VI; SEVERE FORM).
T-> C -> R (IN MPS-VI; SEVERE FORM).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     R -> W (IN MPS-VI; INTERMEDIATE FORM)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /FTIG-VAR_007298.

R -> Q (IN MES-VI; INTERMEDIATE FORM)
/FTIG-VAR_007299.
Y -> C (IN MES-VI; MILD/INTERMEDIATE)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MILD/SEVERE FORM).
                                                                                                                                              MIM; 2/2200;
Interpor, 1PR000917; Sulfatase.
Pfam; PR00884; Sulfatase; 1.
PROSITE; PS00149; SULFATASE_1; 1.
BYGSITE; PS00149; SULFATASE_2; 1.
Bydrolase; Signal; Glycoprotein; Lysosome; Mucopolysaccharidosis; Disease mutation; Polymorphism; 3D-structure.
SIGNAL 1 36 OR 38 (POTENTIAL).
CHAIN 37 533 ARYLSULFATASE B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       H -> P (IN MPS-VI; MILD/SEVERE FORM).
/FTId-VAR_007304.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /FTIG=VAR_007301.
G -> R (IN MPS-VI; SEVERE FORM).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /FIId-VAR_007300.
L -> P (IN MPS-VI; MILD FORM).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C -> Y (IN MPS-VI; MILD FORM).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 10; DB 1; Length 533;
Pred. No. 0.13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Indels
                                                                                                                                                                                                                                                                                  2-AMINO-3-OXOPROPIONIC ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5983FB6911C4789A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               L -> P (IN MPS-VI;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /FTId=VAR_007302.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FTIG=VAR_007303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /FTId=VAR_007306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FTIG=VAR_007305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          V -> M (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100.0%; Prec. ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      535
                                                                                                                                                                                                                                                                   POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                          JOINED.
JOINED.
JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MW.
EMBL, X72737; CAA51272.1; JG
EMBL, X72739; CAA51272.1; JG
EMBL, X72739; CAA51272.1; JG
EMBL, X72740; CAA51272.1; JG
EMBL, X72741; CAA51272.1; JG
EMBL, X72741; CAA51272.1; JG
EMBL; S57777; AAB19988.1; J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          358
59687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 100.
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                  147
91
155
1192
1447
1447
188
229
239
458
458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        137
                                                                                                              04-FEB-98
                                                                                                A35078; A35078
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      171 THMVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       533 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      140 THMVGKWHLG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                091
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       393
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         358
                                                                                                                                                                                                                                                                   147
91
117
121
181
181
405
405
279
279
                                                                                                                                                                                                                                                                                                                                                                                              366
426
458
92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 302
                                                                                                                          253200;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ARSB_FELCA
                                                                                                                1FSU;
                                                                                                                                                                                                                                                                  ACT_SITE
MOD_RES
                                                                                                                                                                                                                                                                                                                         DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                               DISULFID
                                                                                                                                                                                                                                                                                                             DISULFID
                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                         VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /ARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /ARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /ARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /ARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ARSB_FELCA
                                                                                                   PIR;
                                                                                                                PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
```

ARSB

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseélsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PATHWAY: SIXTH STEP IN DE NOVO PURINE BIOSYNTHESIS.
SUBUNIT: HOMOOCTAMER (BY SIMILARITY).
SIMILARITY: TO OTHER BACTERIAL PURE, ALSO TO AIR CARBOXYLASE FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hemicentrotus pulcherrimus (Sea urchin).
Bukaryota, Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea: Euechinoidea; Echinoea; Echinoida; Strongylocentrotidae;
Hemicentrotus;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JAN-1990 (Rel. 13, Created)
01-JAN-1990 (Rel. 13, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
Arylsulfatase precursor (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase)
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sugiura M., Tabata S.; "Sequence analysis of the genome of the unicellular cyanobacterium
                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
MEDLINE-96127529; Pubmed-8590279;
Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CATALYTIC ACTIVITY: 1-(5-phosphoribosyl)-5-amino 4-
imidazolecarboxylate = 1-(5-phosphoribosyl)-5-aminoimidazole
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Purine biosynthesis; Lyase; Decarboxylase; Complete proteome.
SEQUENCE 176 AA; 18603 MW; A1F87A630F18F9E4 CRC64;
                                                                                                01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Phosphoribosylaminoimidazole carboxylase catalytic subunit
                                                                                                                                                                                                                                                                                       Cyanobacteria; Chroccoccales; Synechocystis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 8; DB 1;
Pred. No. 4.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  551 AA
                                                    176 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mismatches
                                                                                                                                                                                                            (EC 4.1.1.21) (AIR carboxylase) (AIRC)
PURE OR SLL0901.
                                                                                                                                                                                                                                                                   Synechocystis sp. (strain PCC 6803)
                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR000031; AIR_carboxyl.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1.5%; 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; D64006; BAA10848.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00731; AIRC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
8; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            248 QILASHNP 255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                131 QILASHNP 138
                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=1148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P09028;
                                                 PUR6_SYNY3
Q55498;
01-NOV-1997 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ARS_HEMPU
P14000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FUNGI.
                                                                                                                                                                                                                                                                                                  Bacteria;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (ARS).
RESULT 5
PUR6_SYNY3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ARS_HEMPU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DATA PART BY A SECOND COLOR CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Genomics 14:403-411(1992).
-1- CATALYTIC ACTIVITY: Hydrolysis of the 4-sulfate groups of the N-acetyl-D-galactosamine 4-sulfate units of chondroitin sulfate and
                           01-FEB-1994 (Rel. 28, Created)
01-FEB-1994 (Rel. 28, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
Arylsulfatase B precursor (EC 3.1.5.12) (ASB) (N-acetylgalactosamine-
                                                                                                                                                                                                                                                                                                                                                                                                                                 "Feline arylsuifatase B (ARSB): isolation and expression of the CDNA, comparison with human ARSB, and gene localization to feline chromosome Al.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -:- SUBUNIT: HOMODIMER.
-:- SUBCELLULAR LOCATION: Lysosomal.
-:- SUBCELLULAR LOCATION: Lysosomal.
-:- DISEBARE: DEFECTS IN ARSB ARE THE CAUSE OF MAROTEAUX-LAMY
SYNDROME; ALSO KNOWN AS MUCOPOLYSACCHARIDOSIS TYPE VI (MPS-VI).
MPS-VI HAS BEEN DESCRIBED IN SIAMESE CATS.
-:- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-93052342; PubMed-1427856;
Jackson C.E., Yuhki N., Desnick R.J., Haskins M.E., O'Brien S.J.,
Schuchman E.H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                         Felis silvestris catus (Cat).
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Carnivora, Fissipedia, Felidae, Felis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hydrolase; Signal; Glycoprotein; Lysosome; Mucopolysaccharidosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 535;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2-AMINO-3-OXOPROPIONIC ACID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     43A527886A9983C4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (BY SIMILARITY).
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
N-LINED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1.8%; Score 10; DB 1;
100.0%; Pred. No. 0.13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL.
ARYLSULFATASE B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. 0.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PROSITE; PS00523; SULFATASE_1; 1. PROSITE; PS00149; SULFATASE_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1.00.08; Pre-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR000917; Sulfatase.
Pfam; PF00884; Sulfatase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        59753 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; S48472; AAB23941.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    535
149
93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        449
190
281
293
428
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; A44475.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     dermatan sulfate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            171 THMVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                123
183
407
190
281
293
428
460
535 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4-sulfatase) (G4S).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
Matches 10; Conserv
                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HSSP; P15848; 1FSU
                                                                                                                                                                                                                                                                         NCBI_TaxID=9685;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    42
149
93
                                                                                                                                                                                                                                                                                                                                                      TISSUE-Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A44475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DISULFID
```

ö

Gaps

ö

NCBI\_TaxID=7650;

CARBOHYD

DISULFID

DISULFID CARBOHYD CARBOHYD CARBOHYD CARBOHYD SEQUENCE Query Match

g

ò

Length 176; 0; Indels S

```
Strongylocentrotus.
            NCBI_TaxID=7668;
                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                             SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NOG1_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                            Yamada K., Akasaka K., Shimada H.;
"Structure of sea-urchin arylsulfatase gene.";
Eur. J. Blochem. 186:405-410(1989).
-!- FUNCTION: MAY BE A STRUCTURAL COMPONENT OF THE EXTRACELLULAR MATRICES INVOLVED IN CELL MOYBENT DURING MORPHOGENESIS.
-!- CATALYTIC ACTIVITY: A phenol sulfate + H(2)0 = a phenol + sulfate.
-!- SUBCELLULAR LOCATION: IN BOTH THE CYTOPLASM AND THE EXTRACELLULAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Strongylocentrotus purpuratus (Purple sea urchin).
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Arylsulfatase precursor (EC 3.1.6.1) (Aryl-sulfate sulphohydrolase)
                                                                 "cDNA cloning, nucleotide sequence and expression of the gene for arylsulfatase in the sea urchin (Hemicentrotus pulcherrimus)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.

N-LINKED (GLCNAC. ..) (POTENTIAL).

N-LINKED (GLCNAC. ..) (POTENTIAL).

""""" (FLCNAC. ..) (POTENTIAL).
                                            Sasaki H., Yamada K., Akasaka H., Suzuki K., Saito A., Sato M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2-AMINO-3-OXOPROPIONIC ACID (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 551;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED (GLCNAC. . .) (P
54ClAAC14D6710C9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                            !ydrolase; Signal; Glycoprotein; Extracellular matrix.
                                                                                                                                                                                                                                     -!- SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1.5%; Score 8; DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
01-0CT-1996 (Rel. 34, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  567 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ARYLSULFATASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0; Mismatches
          SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
                                                                                                                                                                                                                                                                                                                              or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BLOCKED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                               MEDLINE-89030699; PubMed=3181160;
                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF00884; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
                                                                                                                        SEQUENCE FROM N.A.
MEDLINE-90092130; PubMed=2598936;
                                                                                                  Eur. J. Biochem. 177:9-13(1988).
                                                                                                                                                                                                                                                                                                                                                              PIR; S01793; S01793.
PIR; S07089; S07089.
PIRS; P12899; 1AUK.
InterPro; IPR000917; Sulfatase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100.0%;
                                                                                                                                                                                                                                                                                                                                                     EMBL; X17015; CAA34881.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60952
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match
Best Local Similarity 100.،
المراقعة 8: Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21
100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            158
164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     164
213
296
551 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    173 MVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   153 MVGKWHLG 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                21
21
100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            158
                     TISSUE=Pluteus;
                                                                                                                                                                                                                            MATRICES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ARS_STRPU
P50473;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
SEQUENCE
                                                     Shimada
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (ARS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ARS_STRPU
                                                                                                              2
ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseeisb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
STRAIN=S288C / AB972;
STRAIN=S288C / AB972;
MEDLINE=9731371; PubMed=9169875;
Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansorge W., Araujo R., Aparicio A., Barrell B., Badcock K., Benes V., Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M., Chung E.,
                                                                                                                            Dev. Biol. 135:53-65(1989).
-!- CARALYTIC ACTIVITY: A phenol sulfate + H(2)0 = a phenol + sulfate.
-!- SUBCELLULAR LOCATION: Secreted.
-!- SUBCELLULAR LOCATION: EXPRESSION IS CONFINED TO ABORAL ECTODERM
                                                                                                                                                                                                                             CELLS AND THEIR PRECURSORS.

DEVELOPMENTAL STAGE: LOW LEVELS ARE FOUND AT MESENCHYME BLASTULA STAGE (24 HR), LEVELS INCREASE BY LATE GASTRULA STAGE AND ARE MAINTAINED AT PLUTEUS STAGE.

SIMILARITY: BELONGS TO THE SULFATASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NOGI_YEAST STANDARD; PRT; 647 AA.
002892;
001892;
01-MAR-2002 (Rel. 41, Created)
01-MAR-2002 (Rel. 41, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)
Nucleolar GPP-binding protein 1.
NOGI OR YPL093W OR LPG15W.
NOGI OR YPL093W OR LPG15W.
Eukaryota; Fungi; Ascomycota; Saccharomycetales; Saccharomycetales; Saccharomycetales;
                             MEDLINE-89357267; PubMed=2767335;
Yang Q., Angerer L.M., Angerer R.C.;
"Structure and tissue-specific developmental expression of a sea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ARYLSULFATASE.
2-AMINO-3-OXOPROPIONIC ACID (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 567;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       D55B627983A4C4D6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 1;
. 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 8; Dr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1.5%; Scor.
100.0%; Pre
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR000917; Sulfatase.
Pfam; PF00084; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hydrolase; Signal; Glycoprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    62477 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; M28404; AAA30036.1; -. HSSP; P15289; 1AUK.
                                                                                                         urchin arylsulfatase gene.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             173
179
228
542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
8; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           173
179
228
542
567 AA;
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    173 MVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    168 MVGKWHLG 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20
115
```

ó

```
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 10
      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
Churcher C., Coster F., Davis K., Davis R.W., Dietrich F.S.,
Delius H., DiPaolo T., Dubois E., Dusterhoft A., Duncan M., Floeth M.,
Fortin N., Friesen J.D., Fritz C., Goffeau A., Hall J., Hebling U.,
Heumann K., Hilbert H., Hillier L., Hunicke-Smith S., Hyman R.,
Johnston M., Kalman S., Kleine K., Komp C., Kurdi O., Lashkari D.,
Lew H., Lin A., Lin D., Louis E.J., Marathe R., Messenguy F.,
Mewes H.W., Mittipati S., Moestl D., Muller-Auer S., Namath A.,
Nentwich U., Oefner P., Pearson D., Petel F.X., Pohl T.W.,
Purnelle D., Schafer M., Scharfe M., Scherens B., Schramm S.,
Schroeder M., Saicu A.M., Tettelin H., Urrestarazu L.A., Ushinsky S.,
Vierendeels F., Vissers S., Voss H., Walsh S.V., Wambutt R., Wang Y.,
Wedler E., Wedler H., Winnett E., Zhong W.W., Zollner A., Vo D.H.,
                                                                                                                                                                            "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
Nature 387:103-105(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN=5288C;
MEDLINE-55293223; PubMed=7774808;
Hengartner C.J., Thompson C.M., Zhang J., Chao D.M., Liao S.-M.,
Koleske A.J., Okamura S., Young R.A.;
"Association of an activator with an RNA polymerase II holoenzyme.";
Genes Dev. 9:897-910(1995).
                                                                                                                                                                                                                                                                               -i- SUBCELLULAR LOCATION: Nuclear; nucleolar.
-i- SIMILARITY: BELONGS TO THE GTP1 / OBG FAMILY. NOG SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fung1; Ascomycota; Saccharomycotles;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 647;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Indels
                                                                                                                                                                                                              MEDLINE-21062923; PubMed-11112701;
MEDLINE-21062923; PubMed-11112701;
Park J.-H., Jensen B.C., Kifer C.T., Parsons M.;
"A novel nucleolar G-protein conserved in eukaryotes.";
J. Cell Sci. 114:173-186(2001)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GTP (POTENTIAL).
GTP (POTENTIAL).
640324779AE4D716 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-FEB-1995 (Rel. 31, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
01-CT-1996 (Rel. 34, Last annotation update)
Suppressor of RNA polymerase B SRB9 (SCAl protein).
SRB9 OR SCAl OR SSN2 OR YDR443C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1.5%; Score 8; DB 1;
100.0%; Pred. No. 13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GTP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT; 1420 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. 13;
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            74409 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; U43281; AAB68206.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
8; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            647 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     [2]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               454 OALEEEEE 461
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      48 OALEEEEE 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      GTP-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SRB9_YEAST
P38931;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NP_BIND
NP_BIND
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 9
SRB9_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
     ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                   SEQUENCE FROM N.A.
Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Berno A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                 Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
Hunicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
Winant A., Yelton M., Botstein D., Davis R.W.;
Submitted (Aug-1995) to the RMBL/GenBank/DDBJ databases.
-i- FUNCTION: COMPONENT OF THE RNA POLYMERASE II HOLOENZYME AND THE
MEDIATOR OF ACTIVATION SUBCOMPLEX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CBIL OR STM2024.
Salmonella typhimurium.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JUN-1994 (Rel. 29, Created)
01-MAR-2002 (Rel. 41, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)
Precorrin-2 C20-methyltransferase (EC 2.1.1.130) (S-adenosyl-L-methlonine--precorrin-2 methyltransferase) (SP2MT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Roth J.R., Lawrence J.G., Rubenfield M., Kleffer-Higgins S., Church G.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 1420;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7F6CF4BBE0FAC918 CRC64;
Yuryev A., Corden J.L.;
Submitted (APR-1994) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLY-SER.
POLY-LEU.
POLY-GLN.
D -> E (IN REF. 2).
E -> V (IN REF. 2).
T -> S (IN REF. 2).
VK -> GE (IN REF. 2).
T -> P (IN REF. 2).
Y -> S (IN REF. 2).
Y -> S (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A., AND SEQUENCE OF N-TERMINUS
                                                                                                                                                                                                                                                                                                                  -1- SUBCELLULAR LOCATION: Nuclear (Potential).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 1;
. 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            237 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 8; I
Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1.5%; Sc.
100.0%; Pre
0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-93273696; PubMed-8501034;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; U23812; AAA91316.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; U09176; AAA18614.1; -. EMBL; U33007; AAB64875.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                160000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          816
1008
1136
38
812
859
878
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match
Best Local Similarity
کامت 8; Conserve
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PIR; S48538; S48538.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TRANSFAC; T02153; -. SGD; S0002851; SSN2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1420 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1033 LPLDNATL 1040
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      153 LPLDNATL 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Nuclear protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-602;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CBIL_SALTY
Q05593;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Salmonella.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
```

```
HSSP; P25156; 1PMA.
                                                                                                                                                                                                                                                                                                                                                                                                        238 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             specificity.
                                                                                                                                                                                                                                                                                                                                                              28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  51 EEEEEGA 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      81 EEEEEGA 87
                                                                                                                                                                                                                                                                                                                                           Glycoprotein.
CARBOHYD 28
CARBOHYD 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PSA3_ACACA
                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PSA3_ACACA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its muse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                              STRAIN=LT2 / SGSC1412 / ATCC 700720;
MEDLINE=21534948; PubMed=11677609;
McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P., Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D., Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K.;
                                                                                                                                                                                                                                                                                                                                                                                                                        -1- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS SUMT, CYSG, CBIF/COBM AND CBIL/COBI.
                                                                                                                                                                                                                                                                              Nature 413:852-856(2001).
-!- FUNCTION: METHYLATES PRECORRIN-2 AT THE C-20 POSITION TO PRODUCE PRECORRIN-3A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                          'Complete genome sequence of Salmonella enterica serovar Typhimurium
"Characterization of the cobalamin (vitamin B12) biosynthetic genes
                                                                                                                                                                                                                                                                                                                                               s-S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-85160822; PubMed-2984429; McGeoch D.J., Dolan A., Donald S., Rixon F.J.; "Sequence determination and genetic content of the short unique
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                         -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + precorrin-2 adenosyl-L-homocysteine + precorrin-3A.
-1- PATHWAY: COBALAMIN BIOSYNTHESIS.
-1- SUBUNIT: HOMOLIMER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 237;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Herpes simplex virus (type 1 / strain 17).
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Alphaherpesvirinae; Simplexvirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-JAN-1988 (Rel. 06, Created)
01-JAN-1988 (Rel. 06, Last sequence update)
01-OCT-1996 (Rel. 34, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1.3%; Score 7; DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           238 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100.0%; Pred. No. 49; tive 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                  Salmonella typhimurium.";
Bacteriol. 175:3303-3316(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 1.3
Best Local Similarity 100.
Matches 7; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=10299;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     184 ALLAQAG 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                57 ALLAQAG 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Glycoprotein G.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VGLG_HSV11
P06484;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GG OR US4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VGLG_HSV11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ID DTT ACCOOK OCCOOK ACCOOK AC
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@ib-sib.ch).
                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenseélsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Xu P., Zot H.G.;
Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: THE PROTEASOME IS A MULTICATALYTIC PROTEINASE COMPLEX
WHICH IS CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH ARG,
PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT NEUTRAL OR
SLIGHTLY BASIC PH. THE PROTEASOME HAS AN ATP-DEPENDENT PROTEOLYTIC
region in the genome of herpes simplex virus type 1.";
J. Mol. Biol. 181:1-13(1985).
-!- MISCELLANBOUG: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
2: GH, GB, GC, GG, GD, GI, AND GE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -!- PATHWAY: Involved in an ATP/ubiquitin-dependent non-lysosomal proteolytic pathway.
-!- SUBBUXT: THE PROTEASOME IS COMPOSED OF AT LEAST 15 NON IDENTICAL SUBBUXT: THE PROTEASOME IS COMPOSED OF AT LEAST 15 NON IDENTICAL SUBBUXITS WHICH FORM A HIGHLY ORDERED RING-SHAPED STRUCTURE (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIMILARITY).
SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
SIMILARITY: BELONGS TO PEPTIDASE FAMILY TIA; ALSO KNOWN AS THE PROTEASOME A-TYPE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CATALYTIC ACTIVITY: Cleavage at peptide bonds with very broad
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
F7FDC2867E834B92 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15-DEC-1998 (Rel. 37, Created)
15-DEC-1998 (Rel. 37, Last sequence update)
15-DEC-1998 (Rel. 37, Last sequence update)
Proteascome subunit alpha type 3 (EC 3.4.25.1) (Fragment).
Acanthamoeba castellanii (Amoeba).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 238;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 7; DB 1;
Pred. No. 49;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Eukaryota; Acanthamoebidae; Acanthamoeba.
NCBI_TaxID=5755;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1.3%; Scor
100.0%; Pred
0; M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25238 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; L00036; AAA96684.1; -. EMBL; X14112; CAA32281.1; -. EMBL; X02138; CAA26058.1; -. PIR; A05239; QQBE74.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; U85398; AAB41645.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
7; Conserve
```

œ

```
82
DKLAAEG 8
                                                                                                            30-MAY-2000
                                                                                HSLO_SYNY3
P73910:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRUB_BACHD
Q9KA80;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 15
TRUB_BACHD
            97
                                                                   HSLO_SYNY3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
0
                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
STRAIN=MC58 / SEROGROUP B;
STRAIN=MC58 / SEROGROUP B;
MEDLINE=2017575; pubMed=10710307;
Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J., Nelson W.C., GWinn M.L., DeBoy R., Peterson J.D., Hickey E.K., Nelson W.C., GWinn M.L., DeBoy R., Peterson J.D., Hickey E.K., Matt D.H., Salzberg S.L., White O., Fleistchmann R.D., Dougherty B.A., Mason T., Clecko A., Parksey D.S., Blair E., Cittone H., Clark B.B., Cotton M.D., Utterback T.R., Khouri H., Oln H., Vamathevan J., Scalatov V., Masignani V., Pizza M., Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.; "Complete genome sequence of Neisseria meningitidis serogroup B strain
                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                 16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Pantoate--beta-alamine ligase (EC 6.3.2.1) (Pantothenate synthetase)
(Pantoate activating enzyme).
                                                                                                                                                                                                                                                                                                                                                                                                    Neisseria meningitidis (serogroup B).
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -i- CATALYTIC ACTIVITY: ATP + (R)-pantoate + beta-alanine = AMP diphosphate + (R)-pantothenate.
-i- PATHWAY: Pantothenate blosynthesis; last step.
-i- SIMILARITY: BELONGS TO THE PANTOTHENATE SYNTHETASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             .;
0
                                                                                                                                                                 ö
                                                                                                                                    Length 252;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 278;
                                                                                                                                                                 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   pase; 1.
Ligase; Complete proteome.
W; 5E71D92FB0E6F6B3 CRC64;
                                                                  Proteasome, Hydrolase, Protease.
NON_TER 1 1 1
SEQUENCE 252 AA; 27728 MW; 71898276B50C2E79 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 1;
o. 56;
                                                                                                                                   1.3%; Score 7; DB 1;
100.0%; Pred. No. 52;
rative 0; Mismatches
                                                                                                                                                                                                                                                                                          278 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1.3%; Score 7; DB 1
100.0%; Pred. No. 56;
Live 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro: IPR003721; Pantoate_ligase. Pfam; PF02569; Pantoate_ligase; 1. Pantothenate biosynthesis; Ligase; Corsconce 278 AA; 31080 MW; 5E71D9;
                                                                                                                                                                                                                                                                                          PRT;
          InterPro; IPR001353; Proteasome.
InterPro; IPR000426; Proteasome_A.
Pfam; PF00227; proteasome; 1.
PROSITE; PS00388; PROTEASOME_A; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AE002439; AAF41282.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Science 287:1809-1815(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity 100.
                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                    Ouery Match
Best Local Similarity
Matches 7; Conserv
MEROPS; T01.977;
                                                                                                                                                                                                                     239 QALEEEE 245
                                                                                                                                                                                           48 QALEEEE 54
                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=491;
                                                                                                                                                                                                                                                                                                                                                                                       PANC OR NMB087
                                                                                                                                                                                                                                                                                          PANC_NEIMB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local S
Matches 7
                                                                                                                                                                                                                                                                                                         P57036;
                                                                                                                                                                                                                                                              RESULT 13
PANC_NEIMB
                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                           ò
```

104 DKLAAEG 110

ò

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an enail to license@ib-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            THERMALLY UNFOLDING AND OXIDATIVELY DAMAGED PROTEINS FROM IRREVERSIBLE AGGREGATION. PLAYS AN IMPORTANT ROLE IN THE BACTERIAL DEFENSE SYSTEM TOWARD OXIDATIVE STRESS (BY SIMILARITY). SUBCELLULAR LOCATION: CYLOPIDSMIC (By SIMILARITY). PTW: UNDER OXIDIZING CONDITIONS TWO DISULFIDE BONDS ARE FORMED INVOLVING THE REACTIVE CYSTEINES. UNDER REDUCING CONDITIONS ZINC IS BOUND TO THE REACTIVE CYSTEINES AND THE PROTEIN IS INACTIVE (BY
                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDLINE-37061201: PubMed-8905231;
MEDLINE-37061201: PubMed-8905231;
Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
Miyajima N., Hirosawa M., Sugiura M., Sasamoto S., Kimura T.,
Hisouchi T., Mateuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
Tabata S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions "; DNA Res. 3:109-136(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Chaperone; Redox-active center; Zinc; Complete proteome.
1242 REDOX-ACTIVE (BY SIMILARITY).
DISULEID 273 276 REDOX-ACTIVE (BY SIMILARITY).
SEQUENCE 302 AA; 32205 MW; 0EB3448B06E67AAA CRC64;
                                                                                     30-MAY-2000 (Rel. 39, Last sequence update)
10-CCT-2001 (Rel. 40, Last annotation update)
13 kDa chaperonin (Heat shock protein 33 homolog) (HSP33).
HSLO OR SLL1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 302;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                         synechocystis sp. (strain PCC 6803).
Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
NCBI_TaxID=1148;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1.3%; Score 7; DB 1;
100.0%; Pred. No. 61;
ive 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY).
SIMILARITY: BELONGS TO THE HSP33 FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   304 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (Rel. 40, Created)
(Rel. 40, Last sequ
(Rel. 40, Last anno
                                                              39, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; D90910; BAA17974.1; -. Interpro; IPR000397; Hsp33. Pfam; PF01430; HSP33; 1. ProDom; PD014025; Hsp33; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
Matches 7; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            351 GGIRAVG 357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       13 GGIRAVG 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16-OCT-2001
16-OCT-2001
16-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   THE PART
```

```
ERWA pseudouridine synthase B (EC 4.2.1.70) (tRNA pseudouridine 55 DE SYNTHASS) (PSI55 synthase) (Psi55 synthase) (Psi55 synthase) (Uracil DE MIGHOLOUSE) (Uraci
```

Search completed: July 23, 2002, 19:10:27 Job time: 152 sec

54 EEGALLA 60 ||||||| 275 EEGALLA 281

ò

			•

us-09-495-823-7.oli.rspt

```
968100 human cytom 09267 pacthydactyl 06649 aquifex aeo 098812 bacillus ha 097992 aeropyrum p 065435 beet curly 078764 pyxis plani 068084 human cytom 091408 tupaia herp 091408 tupaia herp 09266 homo sapien 091408 tupaia herp 09265 listeria in 090086 agrotis ips P71285 secherichia 050216 pseudomonas 006836 pseudomonas 006836 pseudomonas 0914082 human immun 094mm7 human immun 094mm7 human immun 094mm7 human immun 094107 oryza saziv
                                                                                                                                                                                                                             Q9u933 paramecium
P87027 schizosacch
O14312 schizosacch
Q9buil homo sapien
          human immun
                  29vws6 drosophila
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kawakani T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Olmori Y., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.; "NEDO human cDNA sequencing project."; "NEDO human cDNA sequencing project."; EMBL, RA027201; BAB15689.1; -. EMBL, FA027201; BAB15689.1; -. SEQUENCE 169 AA; 18905 MW; 4583C92FFFCC3380 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STGKSVWLFNITADPYERVDLSNRYPGIVKKLLRRLSQFNKTAVPVRYPPKDPRSNPRLN 547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
         011847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11.5%; Score 63; DB 4; Length 169; 100.0%; Pred. No. 4.7e-57; Live 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                        Q9H5D8;
01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
CDNA: FLA25548 FIS, CLONE LNG08487.
                                                                                                                                                                                                                                                                                                                                  169 AA
                                                                                                                                                                                                                                                                                    ALIGNMENTS
                                                    Q9K8T2
Q9Y992
Q9Y992
Q65435
Q68084
Q968C6
Q96866
Q91TP8
                                                                                                                                     Q9NB86
P71285
O50216
O33496
Q91082
Q91082
Q95RM7
Q94LB7
                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                              Q9U933
P87027
O14312
Q9BUI1
 100
110
110
110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 100.
Matches 63; Conservative
                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ||||
GGV 120
                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE=LUNG;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          548 GGV 550
Query Match
                                                                                                                                                                                                                                                                                                                                  09H5D8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28
                                                                                                                                                                                                                                                                                                               RESULT
Q9H5D8
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9nju7 helix pomat
Q9nju8 helix pomat
Q9vq95 drosophila
Q9hyr2 pseudomonas
Q9bx3 listeria in
Q18924 caenorhabdi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9ckeO pasteurella
Q9v6f2 drosophila
Q25384 hemicentrot
O16138 heliocidari
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q9vvm4 drosophila
Q93d92 streptococc
Q9vvk9 drosophila
Q9axbl oryza sativ
Q92d19 listeria in
                                                            July 23, 2002, 19:06:45; Search time 28.8 Seconds (without alignments) 3303.723 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q9h5d8 homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Description
                                                                                                                   MAPRGCAGHPPPPSPQACVC.....vpvrryppkdprsnprlnggv
                                                                                                                                                                                                    562222
        GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                   Total number of hits satisfying chosen parameters:
                                                                                                                                                               562222 seqs, 172994929 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SUMMARIES
                                                                                                                                                                                                                                                 summaries
                                            OM protein - protein search, using sw model
                                                                                                                                    OLIGO
Gapop 60.0 , Gapext 60.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9NJU8
Q9VJU8
Q9VJU8
Q9HYR2
Q92EX3
Q92EX3
Q9V6F2
Q25384
Q9V6F2
Q953B4
Q9VWM4
Q93D92
Q9AXB1
Q9AXB1
                                                                                                                                                                                                                                                                                                                                                                                                sp_unclassified:*
                                                                                                                                                                                                                                                                                                     sp_human:*
sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                  sp_rodent:*
sp_virus:*
sp_vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                         sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                      sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                               Post-processing: Listing first 45
                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                           sp_archea:*
sp_bacteria:*
                                                                                                                                                                                                                                                                                                                                                                                                                           sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2222222
                                                                                                  US-09-495-823-7
                                                                                                                                                                                                                                                                                                                                                            sp_plant: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB
                                                                                                                                                                                                                                                                  SPTREMBL_19:*
                                                                                                                                                                                                                                                                                           sp_fungi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     266
503
217
217
217
452
453
542
551
559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length
                                                                                                                                                                                                                                                                                                                                  sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query
                                                                                                                                                                                                                                                                                                     63
120
120
88
88
88
88
88
88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score
                                                                                                          Perfect score:
                                                                                                                                     Scoring table:
```

Database :

Word size

Searched:

Sequence:

Run on:

ö

Result ٠ يو

O9NJU7 ~

ACCOCC OCT BEACT B

RESULT

**29NJU7** 

```
REC STRAIN—BERKELEY:

RA Adams M.D. Celniker S.E., Holt R.A., Hoskins R.A., Galle R.F.

Adams M.D. Celniker S.E., Holt R.A., Hoskins R.A., Galle R.F.

RA Adams M.D. Celniker S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.C., Worders Y.-H.C., Blazel R.G., Champe M., Pfelifer B.D.,

RA Brandon R.C., Rogers Y.-H.C., Blazel R.G., Champe M., Pfelifer B.D.,

RA Brandon R.C., Rogers Y.-H.C., Blazel R.G., Champe M., Pfelifer B.D.,

RA Brandon R.C., Rogers Y.-H.C., Blazel R.G., Champe M., Pfelifer B.D.,

RA Ballew R.M., Basu A., Baxenfall G., Nelson C.R., Batlakub D.,

RA Besson K.Y., Bennos P.V., Berman B.P., Bandarl D., Bolshakov S.,

RA Besson K.Y., Bennos P.V., Berman B.P., Brottler P., Bottler P.,

RA Buttis K.C., Busam D.A., Bulke C., Davenport L.B., Davies P.,

RA Cherry J.M., Cawley S., Dablike C., Davenport L.B., Davies P.,

RA Cherry J.M., Cawley S., Dablike C., Davenport L.B., Davies P.,

RA Cherry J.M., Cawley S., Dablike C., Davenport L.B., Davies P.,

RA Cherry J.M., Cawley S., Dablike C., Davenport L.B., Davies P.,

RA Godsen K., Doup L.E., Downess M., Dugan-Focha S., Pleischmann R.,

RA Godsen K., Doup L.E., Downess M., Dugan-Focha S., Pleischmann R.,

RA Godsen K., Gong F., Gorrell J.H., Gu Z., Gaun P., Harris M.,

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,

RA Jalli M., Ralush F., Karpen G.H., Ke Z., Kennison D., Laiz Z.,

RA Jalli M., Ralush F., Karpen G.H., Ke Z., Kennison D., Laiz Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Mattel B. M., Murphy B., Murphy L., Muzsy D. M., Moslson D. L.,

RA Rimmel B.E., Kodira C.D., Kraft C., Kravitz S., Rulp D., Laiz Z.,

Ra Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

Rattorier M., Murphy B., Murphy L., Muzsy D.M., Moslson D.R.,

Reinert M., Murphy B., Murphy L., Wassern D.R., McDier B.,

Rattor R., Remington K., Saunders R.D., Scheel E., Sheal H., Wang X.,

Ra Sher E., Spradling A.C., Stapleton M., Strong R., Sun R.,

Ra Sher E., Spradling A.C., Stapleton W., Shung S., Yao
                                                                                                                                                                                                                                                                           Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                     Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 217;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24439 MW; ECABDBF82A5533C2 CRC64;
                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 8; DB 5;
Pred. No. 12;
0; Mismatches
                                                                                                    217 AA
                                                                                                                                             Created)
                                                                                                                                                                                                                                                Drosophila melanogaster (Fruit fly).
                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001604; Endonuclease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1.5%; Scc.
100.0%; Pre
0; '
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FlyBase; FBgn0031404; CG17006
                                                                                                                                           01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-JUN-2011 (TrEMBLrel. 17, CG17006 PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                    PRELIMINARY;
  82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Matches 8; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SMART; SM00477; NUC; SEQUENCE 217 AA;
                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
71 LENYYVQPICTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   504 ERVDLSNR 511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ERVDLSNR 42
                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-BERKELEY
                                                                                                                                                                                                                                                                     Eukaryota;
                                                                                                                                                                                                                          cc17006
                                                                                                  Q9VQ95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         32
                                                            RESULT
Q9VQ95
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ô
                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. MEDIAG-10772344; MEDIAGE B.A.; MAILSELOCK U., Fischer M., Svendsen I., Halkier B.A.; "Cloning and characterization of two cDNAs encoding sulfatases in the
                                                                                                                                                                                                   Helix pomatia (Roman snail) (Edible snail).
Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
Helicacea; Helicidae; Helix.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Helix pomatia (Roman snail) (Edible snail).
Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
Helicacea; Helicidae; Helix.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-20233298; PubMed-10772344;
Wittstock U., Fischer M., Svendsen I., Halkier B.A.;
"Cloning and characterization of two cDNAs encoding sulfatases in the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 266;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2.2%; Score 12; DB 5; Length 503; 100.0%; Pred. No. 0.0017; ve 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR000917; Sulfatase.
Pfam. PF00884; Sulfatase; 1.
PROSITE; PS00149; SULFATASE_2; 1.
SEQUENCE 503 AA; 55821 MW; AC45334BA694413D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       266 AA; 30391 MW; 2F877BB1737A9C49 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-0CT-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                                                                     Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3.6%; Score 20; DB 5; L6
100.0%; Pred. No. 4.6e-12;
tive 0; Mismatches 0;
                                                            266 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Local Similarity 100.0%; Pred. No. 0.0
Matches 12; Conservative 0; Mismatches
                                                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Interpro; IPR000917; Sulfatase.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                       Roman snail, Helix pomatia.";
IUBMB Life 49:71-76(2000).
EMBL; AFI09925; AAF30403.1; -
HSSP; P15848; 1FSU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Roman snail, Helix pomatia.";
IUBMB Life 49:71-76(2000).
EMBL; AF109924; AAF30402.1; -
HSSP; P15848; IFSU.
                                                                                              (TrEMBLrel, 15, C
(TrEMBLrel, 15, L
(TrEMBLrel, 19, L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   103 LDKLAAEGVKLENYYVQPIC 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        55 LDKLAAEGVKLENYYVQPIC 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                            SULFATASE 2 (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    113 LENYYVQPICTP 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-6536;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=6536;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Simi
Matches 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SULFATASE 1.
                                                                                                01-OCT-2000
                                                                                                                   01-OCT-2000
01-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                             , 7UUN60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09NJUB;
```

**90000** 

ò

m

RESULT Q9NJU8

ò g ö

Gaps

ó

Gaps

```
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoldea; Rhabditidae; Peloderinae; Caenorhabditis.
NCBL_TaxID=6239;
                                                                                                                                                                                                                                                                                               ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium."; Science 282:2012-2018(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
0
                                                                                                                                                                                                                                    DB 16; Length 381;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 452;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0; Indels
                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
STRAIN=BRLSTOL N2;
Du Z., Leimbac D.;
"The sequence of C. elegans cosmid D1014.";
Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Direct Submission.";
Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
EMBL: U53180; AAA96290.1; --
HSSP; P15289; LAUK.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hypothetical protein.
SEQUENCE 452 AA; 51685 MW; 14999D3C1F7891D2 CRC64;
                                                                                                                                              381 AA; 43625 MW; D74B4FEE1C1A720C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-DEC-2001 (TEMBLrel. 19, Last annotation update)
HYPOTHERICAL 51.7 KDA PROTEIN.
                                                                                                                                                                                                                                                                                               ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1.5%; Score 8; DB 5;
"Comparative genomics of Listeria species.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        453 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      452 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 100.0%; Pred. No. ...
Matches 8; Conservative 0; Mismatches
                                                                                                                                                                                                                                    1.5%; Score 8; DB 1
100.0%; Pred. No. 20;
Live 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=BRISTOL N2;
MEDLINE=99069613; Pubmed=9851916;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROMITE: PS0084; SULFATASE: 1.
PROSITE: PS00523; SULFATASE_1; 1.
PROSITE: PS00149; SULFATASE_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR000917; Sulfatase.
                           Science 294:849-852(2001).
EMBL; AL596164; CAC95568.1;
ListList; LIN00335; --
Complete proteome.
SEQUENCE 381 AA; 43625 MW
                                                                                                                                                                                                                                    Ouery Match 1.5
Best Local Similarity 100.
Matches 8; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Caenorhabditis elegans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A. STRAIN-BRISTOL N2;
                                                                                                                                                                                                                                                                                                                                                                                         181 GEYNNSII 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         173 MVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           134 MVGKWHLG 141
                                                                                                                                                                                                                                                                                                                                                          315 GFYNNSII 322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Waterston R.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q18924
Q18924;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9CKEO
Q9CKEO;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      D1014.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
Q9CKE0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AC AC
   RT
DR
DR
SQ
                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAC DDT THE REPORT OF THE REPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C STRAIN-CLIP 1152 / SEROVAR 6A;

C STRAIN-CLIP 1152 / SEROVAR 6A;

R Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,

A Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,

Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,

Charbit A., Chetouani F., Couve E., de Daruvar A., Denoux P.,

Bomian E., Dominguez Bernal G., Duchaud E., Durant L., Dussurget O.,

A Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,

A Gautler L., Goebel W., Gomez-Lopez N., Hauf T., Hauf J., Jackson D.,

A Jones L.-M., Raerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,

A Madueno E., Mattournam A., Mata Vicente J., Ng E., Nedjari H.,

A Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,

A Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,

Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN=ATCC 15692 / PAO1;
MEDLINE=20437337; PubMed=10984043;
StOVET C.K., Pham X. O.T., Erwin A.L., Mizoguchi S.D., Warrener D.,
Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
Reizer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.;
"Complete genome sequence of Pseudomonas aeruginosa PAO1, an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
0
                                                                                                                                                                                                                                                                                                                                                          Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1.5%; Score 8; DB 16; Length 330; 100.0%; Pred. No. 18; 0; Indels ative 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        330 AA; 34003 MW; 6F032604DF695F1E CRC64;
                                                                                                                                                                      01-MAR-2001 (TrEMBLrel. 16, Created)
1-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
3-OXOACYL-ACYL-CARRIER-PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Interpro; IPPN02106; AA_tRNA_ligase_II.
PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_I.
Complete proteome.
SEQUENCE 330 AA; 34003 MW; 6F032604DF695F1F C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Listeria.
NCBL_TaxID=1642;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             381 AA.
                                                                                                                     330 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; AE004755; AAG06721.1; -. HSSP; P24249; 1EBL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, LIN0335 PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 100..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            opportunistic pathogen."; Nature 406:959-964(2000).
                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                              Pseudomonas aeruginosa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     171 EEEGALLA 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Listeria innocua.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 53 EEEGALLA 60
                                                                                                                                                                                                                                                                                            FABH2 OR PA3333
                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=287;
                                                                                                                                                                                                                                                                                                                                                                                      Pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q92EX3
Q92EX3;
                                                                                                               Q9HYR2
                                                                                                                                              Q9HYR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ø
                                                       RESULT
Q9HYR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              092EX3
```

q

ó

ö

Gaps

```
01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-DEC-2001 (TrEMBLrel. 19,
                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 100....
8; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 1-105 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Submitted (NOV-1995)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                              120 PICTPSRS 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                              11111111
71 PICTPSRS 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hemicentrotus.
NCBI_TaxID=7650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ARYLSULFATASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TISSUE-SPERM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUE-SPERM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TISSUE-SPERM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TISSUE-SPERM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Yamada K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Shimada
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                025384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 10
Q25384
  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Adamstides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandrell M.D., Zhang O., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfelifer B.D.,
Ann K.H., Doyle C., Baxerdale J., Bayraktaroglu L., Beasley E.M.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Besson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Andreys D., Botchan M.R., Bouck J., Erokstein P., Botchar A., Dahlke C., Davenport L.B., Davis S.M.,
Cablery J.M., Cawley S., Dahlke C., Davenport L.B., Davis S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
A Dotson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C.,
Cabrielian A.E., Garraz C., Ferriacr S., Fleischmann W.,
R. Fosler C., Gabriellian A.E., Garraz C., Ferriacr S., Fleischmann M.,
Antris N.L., Harrey D., Hehnan T.J., Hernandez J.R., Houst,
A Harris N.L., Harrey D., Hehnan T.J., Wei M.-H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H. Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai z.,
                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                   May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.; "Complete genomic sequence of Pasteurella multocida Pm70."; Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).

EMBL, AE006204; AAK03766.1; -. 98:3460-3465(2001).

InterPro: IPR000056; Lipocln_cytFABP.

InterPro: IPR000917; Sulfatase.

PROSTEE; PS00213; LIPCCALIN; UNKNOWN_1.

PHYPOCHELICAL protein; Complete proteome.

SEQUENCE 453 AA; 51449 MW; 83B91532C6252B0E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                    Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Pasteurella.
                                                                                                                                                                                                                                                                                                                                                            Length 453;
                                                                                                                                                                                                                                                                                                                                                                                       Indels
01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
HYPOTHETICAL PROTEIN PM1682.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
CG8646 PROTEIN.
                                                                                                                                                                                                                                                                                                                                                           1.5%; Score 8; DB 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   542 AA.
                                                                                                                                                                                                                                                                                                                                                                           Pred. No. 23;
                                                                                                                                                                                                                                                                                                                                                                    100.0%; Pred. ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                          STRAIN-PM70;
MEDLINE-21145866; PubMed-11248100;
                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 100.
Matches 8; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                       Pasteurella multocida.
                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                  100 TPTLDKLA 107
                                                                                                                                                                                                                                                                                                                                                                                                                                   1111111
29 TPTLDKLA 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-7227;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-BERKELEY;
                                                                                                               NCBI_TaxID=747;
01-JUN-2001 (
01-JUN-2001 (
01-DEC-2001 (
                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09V6F2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09V6F2
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
 셤
```

```
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
A Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
A Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshreft A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
A Molson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
A Ralazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
A Ralazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
A Ralacreft K., Saunders R.D.C., Scheeler F., Shen H.,
A Shue B.C., Siden Kiamos I., Simpson M., Skupski M.P., Smith T.,
A Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Syirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
A Mang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
A Mang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
A Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
A Zheng X.H., Myers E.W., Rubin G.M., Venter J.C.,
Science 287:1285-2195(2000).
B The genome sequence of Drosophila melanogaster.";
B RBBL, Ab003821, AAF58475.1; -.
B RBBL, Absoe; FBH0033753 (CG6466.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Akasaka K., Sakamoto N., Yamamoto T., Morokuma J., Fujikawa N.,
Takata K., Eguchi S., Shimada H.;
"Corrected structure of the 5'flanking region of arylsulfatase gene of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hemicentrotus pulcherrimus (Sea urchin).
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 542;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Submitted (SEP-1989) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Interpro; IPR000917; Sulfatase.
Pfam; PF00884; Sulfatase: 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
SEQUENCE 542 AA; 60278 NW; EACD9C14DD2B3204 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Yamada K., Akasaka K., Shimada H.;
"Structure of sea-Urchin arylsulfatase gene.";
Eur. J. Biochem. 186:405-410(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 8; DB 5; Pred. No. 27; 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1.5%; Scor-
100.0%; Pred
0; M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=90092130; PubMed=2598936;
```

S

õ

```
SUCENCE FROM N.A.

RAY STRAIN-BERKELER:

RAY MEDLINE-20196006; PubMed=10731132;

RADADARLE S.C. Holt R.A., Evans C.A., Gocayne J.D.,

RADADARLE S.C., Foller S.E., Lip P.W., Hoskins R.A., Galle R.F.,

RAM Amanatides P.G., Scherer S.E., Lip P.W., Hoskins R.A., Galle R.F.,

RAM Amanatides P.G., Scherer S.E., Lip P.W., Hoskins R.A., Galle R.F.,

RAM Brandon G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,

RAM Enandon G.G., Mortman J.R., Yandell M.D., Zhang O., Chen L.X.,

RAM Abril J.F., Agbayani A., An Helt G., Nelson C.R., Miklos G.L.G.,

RAM Abril J.F., Agbayani A., An H.J., Andrews-Pfennach, C., Backlain D.,

Ballew R.M., Basu A., Baxendale J., Bayraktaroju L., Beasley E.M.,

RADIS B.J., Bencos P. W., Berman B.P., Bhandari D., Bolshakov S.,

RAM Beeson K.Y., Benos P.W., Berman B.P., Bhandari D., Bolshakov S.,

RAM Grerry J.M., Cawley S., Dalike C., Daven P., Brottler P., Enctler P.,

Buttls K.C., Busam D.A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RAM Grerry J.M., Cawley S., Dalike C., Down P., Borties P.,

RAM Gross B., Delcher A., Deng Z., Gang N.S., Gelbart W.M., Glasser K.,

RAM Godson K.J., Evangelista C.C., Ferriaz C., Ferriac S., Fleischmann W.,

RAM Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,

Alalali M., Kalush F., Karpen G. H., Ke Z., Gang N.S., Globart W.M., Match B. McIncosh T.C., Mocleod W.P., Morky D.M., Nathol M. M.,

RAM Harris N.L., Match E. McIncosh T.C., Mocleod W.P., Morlerson D.L.,

RAM Mount S.M., Wodlyw M., Wurphy B., Murphy L., Muzny D.M., Nathol M.,

RAM B. Nelson K.A., Howlinstock G.M., Welson D.L.,

RAM Reinert K., Remington K., Standers R., Penter E., Shen H.,

RAM Sarraman D.A., Weinstock G.M., Weinstock J., Puri, V.,

RAM R., Mang Z.-Y., Wassarman D.A., Weinstock G.M., Weinstock J., Weinstock J., Stabo G., Zhen G., Zhen M., Wang Z.-Y., Wang S., Zako G., Zhen G., Stabo G., Zhen G., Stabe G., Zhen M., Wang Z., Zhen K., Weinstock G.M., Weinstock S., Zhen G., Stabo G., S
                                             Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pteryyota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
NCBL_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 579;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               65053 MW; EBF2F4F13D759298 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 1.5%; Score 8; DB 5;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
                            Drosophila melanogaster (Fruit fly).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PROSITE; PS00149; SULFATASE_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FlyBase; FBgn0036768; CG7402.
InterPro; IPR000917; Sulfatase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (TrEMBLrel. 19, C
(TrEMBLrel. 19, I
(TrEMBLrel. 19, I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF00884; Sulfatase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       127 VGKWHLGF 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               174 VGKWHLGF 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-DEC-2001 (
01-DEC-2001 (
01-DEC-2001 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q93D92
Q93D92;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         093092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οg
        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DATA TA
                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-98298333; PubMed-9634485;
Haag E.S., Raff R.A.;
"Isolation and characterization of three mRNAs enriched in embryos of
the direct-developing sea urchin heliocidaris erythrogramma: evolution
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota; Metažoa; Ēchinodērmata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida; Echinometridae;
Heliocidaris.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ó
                                                                                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                                                                                       Length 551;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 559;
                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61915 MW; 13DE7DDCEBD05752 CRC64;
                                                                                                                                                                                                                                                                            60943 MW; 544315A4880E12C1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0;
the sea urchin, Hemicentrotus pulcherrimus.";
Dev. Growth Differ. 36:633-636(1994).
EMBL; X16679; CAR34667.1;
EMBL; X16680; CAR34667.1; JOINED.
EMBL; X16681; CAR34667.1; JOINED.
EMBL; X16681; CAR34667.1; JOINED.
EMBL; X16681; CAR34667.1; JOINED.
EMBL; X16683; CAR34667.1; JOINED.
EMBL; X16689; CAR34667.1; JOINED.
EMBL; X16689; DAN34667.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 5;
o. 28;
                                                                                                                                                                                                                                                                                                                                    Ouery Match 1.5%; Score 8; DB 5; Best Local Similarity 100.0%; Pred. No. 28; Matches 8; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1.5%; Score 8; DB 5
100.0%; Pred. No. 28;
ive 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Heliocidaris erythrogramma (Sea urchin).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Dev. Genes Evol. 208:188-204(1998).
EMBL; AF013158; AAC27821.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HSSP, P15289, IAUK.
InterPro; IPR000917; Sulfatase.
InterPro; IPR000917; Sulfatase.
PROSITE; PS00523; SULFATASE_1; IPROSITE; PS00149; SULFATASE_2; I
                                                                                                                                                                                                            Pfam; PF00884; Sulfatase; 1.
PROSITE; PS00523; SULFATASE_1; 1.
PROSITE; PS00149; SULFATASE_2; 1.
                                                                                                                                                                                         IPR000917; Sulfatase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-JAN-1998 (TrEMBLrel. 05, 01-JAN-1998 (TrEMBLrel. 05, 01-DEC-2001 (TrEMBLrel. 19, ARYLSULFATASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MAY 2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-JUN-2001 (TrEMBLrel. 17, CG7402 PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        of larval ectoderm.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
Matches 8; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                         173 MVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                    153 MVGKWHLG 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       173 MVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |||||||||
| 159 MVGKWHLG 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=7634;
                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             09VVM4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               016138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         016138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09VVM4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 11
```

ö

090VM4 ID 099 AC 099 DT 011 DT 011 DE CC

g

ò

```
01-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                  Q9AXB1
                                                                                                                                                                                                                                                                                                                                                                                      RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                     Q9AXB1
 δ
                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Amanatides P.C., Scherer S.E., Holf R.A., Evans C.A., Gocayne J.D.,
Amanatides P.C., Scherer S.E., Holf R.A., Hoskins R.A., Galle R.F.,
A Adams M.D., Celliker S.E., Holf W., Hoskins R.A., Galle R.F.,
A Goorge R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Button G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
An K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Brotsteir P.,
Besson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
A Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
A Burtis K.C., Busam D.A., Butler C., Davenport L.B., Davies P.,
A Burtis K.C., Busam D.A., Butler C., Davenport L.B., Davies P.,
A Burtis K.C., Busam D.A., Butler C., Davenport L.B., Davies P.,
A Burtis K.C., Gabriellan A.E., Gorge C., Ferriera S., Fleischmann W.,
A Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Burtis N.L., Harvey D., Heinan T.J., Hernandez J.R., Houck J.,
A Jalali M., Kalush F., Karpen G.H., Ke Z., Guan P., Harris M.,
A Jalali M., Kalush F., Karpen G.H., Ke Z., Kunison J.A., Ketchum K.A.,
A Jalali M., Mullshina N.V., Mobarry C., Morris J., Moshrefi A.,
A Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
A Merkulov G., Milshina N.V., Mobarry C., Scheeler F., Shen H.,
A Bhazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Rese M.G.,
Rhue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Drosophila melanogaster (Fruit fly).

Bukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:

Pterygota: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha:

Ephydroidea: Drosophilidae: Drosophila.
                                                                                                                                                                                                                                                                         Gaps
                          Streptococcus mutans.
Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
                                                                                                                             lao L., Tanzer J.M.;
Novel sucrose-dependent adhesion cofactors (sdc) in Streptococcus
                                                                                                                                                                                                                                                                       ;
                                                                                                                                                                                                                                          Length 898;
                                                                                                                                                                                                                                                                       0; Indels
                                                                                                                                                                    Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF397166; AAL04089.1; -
SEQUENCE 898 AA; 101903 MW; A2D9CA6F9357AF11 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                          DB 2;
                                                                                                                                                                                                                                                                       Mismatches
                                                                                                                                                                                                                                           Score 8; I
Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-20196006; PubMed-10731132;
                                                                                                                                                                                                                              1.5%; Scor.
100.0%; Pred
                                                                                                                                                                                                                                                                                                                                                                                                                             01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-JUN-2001 (TrEMBLrel. 17,
                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                 195 TFFGSLLG 202
                                                                     NCBI_TaxID=1309;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-BERKELEY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CG5584 PROTEIN.
                                                        Streptococcus
                                                                                                               STRAIN-LT11;
                                                                                                                                                                                                                                                                       ..
8
                                                                                                                                                          mutans.'
                                                                                                                                                                                                                                                                                                                                                                                                                  09VVK9;
                                                                                                                                                                                                                                                                                                                                                                                                  09VVK9
                                                                                                                                                                                                                                                                                                                                                                         RESULT 14
                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                     ô
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner R., Venter E., Wang A. H., Wang X., Wang S., Tector C., Turner R., Venter E., Wang A. H., Wang X., Wang S., Williams S. M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng F.N., Zhong F.N., Zhong F.N., Zhun M., Zhun S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; The genome sequence of Drosophila melanogaster."; Science 287:12185-2195(2000).

HSSP: P15848: IFSU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,
Spermatophyta, Magnollophyta, Liliopsida, Poales, Poaceae,
Ehrhartoldeae, Oryzeae, Oryza.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.

SEQUENCE FROM N.A.

STARIN=CV. NIPPONBARE;

STARIN=CV. NIPPONBARE;

SASAKI T., MACENIMOTO T., Yamamoto K.;

CONCERNO A SACTIVA NIPPONBARE (GA3) GENOMIC DNA, CHROMOSOME 1, PAC Clone:P0501G01.";

CONCERNO SUDMITCH (JUL-2000) to the EMBL/GenBank/DDBJ databases.

REMBL; AP002819; BAB21077.1;

RINTERPRO; IPR001871; DZIP.

RINTERPRO; IPR001871; DZIP.

RINTERPRO; IPR001871; DZIP.

REMBL; PF00534; Glycos_transf_1.

REMBL; PF00534; Glycos_transf_1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1.5%; Score 8; DB 10; Length 1002; 100.0%; Pred. No. 47; ve 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Indels
                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR000917; Sulfatase.
Pfam; PF00884; Sulfatase; 1.
PROSTIE; PO0149; SULFATASE, 2; 1.
SEQUENCE 996 AA; 113514 WW; C92AA3788D082FB3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 1.5%; Score 8; DB 5; Best Local Similarity 100.0%; Pred. No. 47; Matches 8; Conservative 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 1002 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Search completed: July 23, 2002, 19:09:59 Job time: 194 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1.5¢,
100.0%; Pre-
0;
                                                                                                                                                                                                                                                                                                                                   FlyBase; FBgn0036755; CG5584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
Best Local Similarity 100.v
Lac 8; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-JUN-2001 (TrEMBLrel. 01-JUN-2001 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    P0501G01.6 PROTEIN.
P0501G01.6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Oryza sativa (Rice)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 174 VGKWHLGF 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        126 VGKWHLGF 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8 GHPPPPSP 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=4530;
```

9:-

. .

Yeast transcriptionum cytomogalovi Human cytomogalovi Huran cytomogalovi Huran cytomegalovi Huran Taranscriptide HPV E6 protein inh E2 peptide E2-15, Human polypeptide Human EST encoded Drosophila melanog

Drosophila melanog Sea urchin sulfata Drosophila melanog Drosophila melanog

Pseudomonas aerugi

Human brain express Human bone marrow Peptide #108 encod Peptide #100 encod Peptide #100 encod Novel human secret HIV-1 Tat peptide Human polypeptide Novel human secret Human polypeptide Novel human secret Human polypeptide

Novel human secret Novel human diagno Human polypeptide Arabidopsis thalia

Novel human sulfatase polypeptides useful for treating and diagnosing sulfatase-related disorders such as cerebrovascular diseases, acute meningitis, multiple sclerosis, degenerative diseases and tumor

Novel human enzyme Human 25278 sulfat

Human sulfatase pr Drosophila melanog

Human PRO polypept

70.0 52.5 32.5 32.9 32.9 32.5 1.8 1.8

385 289 196 196 181 181 179 109 109 10

Human Human

Query Match 1

Score

Result No.

protein segu

Human | Novel |

WPI; 2001-476214/51. N-PSDB; AAH46863, AAH46867.

Tsai

Human cardiovascul Human peptide #107 Peptide #111 encod Protein #100 encod

0

Word size :

Searched:

Database :

Title: Perfect score:

Sequence:

protein

ĕ

Run on:

Scoring table:

Human immune/haema

```
Sulfatase; 23553 sulfatase; nootropic; neuroprotective; antibacterial; antiinflammatory; vasotropic; antitumor; gene therapy; human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Rudolph-Owen LA,
                                                                                                                                                                                                                                                                                                                                                                                                                                       ALIGNMENTS
                                             ABB61654
AAW13826
AAB86555
AAB86555
AAB86555
AAB86557
AAA13860
AAA13380
AAA3380
AAM2329
AAM2320
AAM2220
AAM2220
AAM2340
AAM5812
AAM5812
AAM58134
AAM65812
AAM58134
AAM65812
AAM3674
AAM65812
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Human 23553 sulfatase polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB85483 standard; Protein; 599
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Glucksmann MA, Williamson M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (MILL-) MILLENNIUM PHARM INC
31-JAN-2001; 2001WO-US03266.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 31-JAN-2000; 2000US-0495823.
WO200155411-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                02-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB85483;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB85483
July 23, 2002, 19:02:54; Search time 38.31 Seconds (without alignments) 1594.638 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              novel protei
PRO708 prote
PRO708 (UNQ3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human 23553 sulfat
Human novel protei
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                             A_Geneseq_032802:*

| SIDSI/gcddata/geneseq/geneseqp-embl/AA1980.DAT:*
| SIDSI/gcddata/geneseqy-embl/AA1981.DAT:*
| SIDSI/gcddata/geneseqy-embl/AA1990.DAT:*
| SIDSI/gcddata/geneseqy-embl/AA1990.DAT:*
| SIDSI/gcddata/geneseqy-embl/AA1990.DAT:*
| SIDSI/gcddata/geneseqy-embl/AA1991.DAT:*
| SIDSI/gcddata/geneseqy-embl/AA1999.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Description
                                                                                                                                                     MAPRGCAGHPPPPSPQACVC......VPVRYPPKDPRSNPRLNGGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Human
                                                                                                                                                                                                                                                                747574
        GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                              Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                               747574 seqs, 111073796 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUMMARIES
                                                                                                                                                                                                                                                                                                                         Post-processing: Listing first 45 summaries
                                                        - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAU14421
AAV14185
AAY41701
AAB44257
AAU29061
AAW25415
AAW25415
AAW25415
AAW25415
AAW25415
AAW25415
AAW25418
AAW25418
AAW864184
AAW866123
                                                                                                                                                                            OLIGO
Gapop 60.0 , Gapext 60.0
                                                                                                                                                                                                                                                                                      Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-495-823-7
550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  289
366
515
515
515
515
1128
569
533
```

```
WO200155437-A2
                                                   Homo sapiens
                                                                                          02-AUG-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     289;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                 ö
                                               cerebrovascular diseases, infections such as acute meningitis, demyelinating diseases including multiple sclerosis, degenerative diseases affecting the cerebral cortex including Alzheimer's disease and Pick disease, spinocerebellar degenerations including spinocerebellar ataxias including Friedreich ataxia, and ataxia telangiectasia, degenerative diseases affecting motor neurons including amyotrophic lateral sclerosis, inborn errors of metabolism such as leukodystrophics, toxic and acquired metabolic diseases, including vitamin deficiencies, and neurocutaneous syndromes (phakamatoses) including neurofibromotosis. The present sequence represents a human 23553 sulfatase polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic;
anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
                              polypeptides
                                                                                                                                                                                                                                                                            ICTPSRSQFITGKYQIHTGLOHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHWVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NIDPIYTKAKNGSWAAGYGIWNTAIQSAIRVQHWKLLIGNPGYSDWVPPQSFSNLGPNRW 480
                                                                                                                                                                                                                                                                                                                                          ictpsrsqfitgkyqihtglqhsiirptqpnclpldnatlpqklkevgysthmvgkwhlg 180
                                                                                                                                                                                                                                                                                                                                                                                                      MYTQRVQQILASHNPTKPIFLYIAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCL 300
                                                                                                                                                                                                                                      1 MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLA 60
                    invention provides 22438, 23553, 25278 or 26212 human sulfatase
                                                                                                                                                                                                                                               FYRKECMPTRRGFDTFFGSLLGSGDYYTHYRCDSPGMCGYDLYENDNAAWDYDNGIYSTQ
                                                                                                                                                                                                                                                                                                                                                                                                                                               DEAINNVTLALKTYGFYNNSIIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SPLLKNKGTVCKEPVHITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Human; novel protein; Antianaemic; osteopathic; antiinflammatory;
                                                                                                                                                                                                                  .;
0
                             polypeptides and polynuclectides. The sulfatase genes and poleare useful for treating disorders involving the brain such as
                                                                                                                                                                                             Length 599;
                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                 1;
                                                                                                                                                                                               DB 22;
                                                                                                                                                                                             Score 385; DB Pred. No. 0; 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAU14421 standard; Protein; 289
Claim 9; Fig 15; 180pp; English
                                                                                                                                                                                              70.0%;
99.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protein #292.
                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                       Best_Local Similarity
Matches 485; Conserv
                                                                                                                                                                599 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HNERIT 486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hnerit 486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    24-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Human novel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAU14421;
                                                                                                                                                                 Sequence
                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ~
                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                                                                      241
                                                                                                                                                                                                                                                                                                                                                                                                                           241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          481
                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                             301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAU14421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
The invention relates to polynucleotides encoding novel human proteins or their active domains. The polypeptides, polynucleotides and antibodies raised against the polypeptides are used in a method of the reatment of a mammal and prevention of disorders caused by the aberrant protein expression or activity. The polypeptides can be used as molecular weight markers, food supplements, and in antibody production. The polypeptides are used to identify compounds which bind to the polypeptides. Polynucleotides of the invention are used as probes and primers, for sequencing, for chromosome or gene mapping, in the production of recombinant proteins, and in generating anti-sense DNA or RNA and in gene therapy. Polypeptides of the invention can be used to target drugs to a tumour, in assays to determine biological activity, to raise antibodies/elicit an immune response, to determine quantitative protein levels, as tissue markers, and to isolate receptors or ligands. Polypeptides of the invention may also be useful in treating platelet disorders, treatened hasorders, regenerating borne, cartilage, tendon, ilyament and/or nerve tissue, wound healing, treating burns, promoting the proliferation, differentiation and survival of stem cells, as a contraceptive, treating osteoporosis and osteoarthitis, anaemia, hisheimer's, parkinson's and Huntington's diseases, amylotrophic lateral sclerosis, stroke, immune deficiencies resulting from bacterial, viral or fungal infection or from autoimmunity, cancer, allergy, asthma, graft-versus-host diseases, nervous system diseases, and infection.

The present sequence represents a protein of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic; thrombolytic; immunogen; antibody; gene therapy; neurological disorder; Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GDYYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQILASHNPTKPIFLYI 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DDOGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITGKYQIHTGLQHS 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Isolated polypeptides useful for treating anti-inflammatory diseases, nervous system disorders, and for regenerating bone and cartilage -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24 MLAMGALAGFWILCLLTYGYLSWGQALEEEEBGALLAQAGEKLEPSTTSTSQPHLIFILA 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDTFFGSLLGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   52.5%; Score 289; DB 22; Length 289;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. 2.4e-285;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .;
O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Example 4; Page 809-810; 894pp; English.
                                                                                                                                 tissue regeneration; immune disorder.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Drmanac RT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1.00.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25-JAN-2001; 2001WO-US02623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25-JAN-2000; 2000US-0491404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2001-451939/48.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               289 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (HYSE-) HYSEQ INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Tang YT, Liu C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-PSDB; AAS22726
```

m

```
98US-0080328
                 366 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     17-MAR-1998;
20-MAR-1998;
20-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20-MAR-1998;
20-MAR-1998;
25-MAR-1998;
26-MAR-1998;
27-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO9946281-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              07-DEC-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           27-MAR-1998;
27-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         27-MAR-1998;
27-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       30-MAR-1998;
30-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       31-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    31-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    31-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY41701;
                                                              Query Match
                   Sequence
                                                                                                                         16
                                                                                                                                                                                    92
                                                                                                                                                                                                                                                                                                                                                                                                     AAY41701
                                                                                                                                                                                                                                                                                                                                                                                      RESULT
    XX
SO
                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                     g
                                                                                                                                                                                    ò
                                                                                                                                                                                                               q
                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                        Human; novel protein; Antianaemic; osteopathic; antiinflammatory; immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic; anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral; antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic; thrombolytic; lammunogen; antibody; gene therapy; neurological disorder; Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis; tissue regeneration; immune disorder.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sclerosis, stroke, immune deficiencies resulting from bacterial, viral or fungal infection or from autoimmunity, cancer, allergy, asthma; graft-versus-host disease, eczema, haemophilia, thrombosis, and infection. anti-inflammatory diseases, nervous system disorders, and infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The invention relates to polynucleotides encoding novel human proteins or their active domains. The polypeptides, polynucleotides and antibodies raised against the polypeptides are used in a method of treatment of a mammal and prevention of disorders caused by the aberrant protein expression or activity. The polypeptides can be used as molecular weight markers, food supplements, and in antibody production. The polypeptides are used to identify compounds which bind to the polypeptides. Polynucleotides of the invention are used as probes and primers, for sequencing, for chromosome or gene mapping, in the production of recombinant proteins, and in generating anti-sense DNA or RNA and in gene therapy. Polypeptides of the invention can be used to target during to a tumour, in assays to determine biological activity, to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      raise antibodies/elicit an immore response, to determine quantitative protein levels, as tissue markers, and to isolate receptors or ligands. Polypeptides of the invention may also be useful in treating platelet disorders, stem cell disorders, regenerating bone, cartilage, tendon, ligament and/or nerve tissue, wound healing, treating burns, promoting the proliferation, differentiation and survival of stem cells, as a contraceptive, treating osteoporosis and osteoarthritis, anaemia, Altheimer's, Parkinson's and Huntington's diseases, amylotrophic lateral
Isolated polypeptides useful for treating anti-inflammatory diseases, nervous system disorders, and for regenerating bone and cartilage -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 The present sequence represents a protein of the invention.
                                                              241 aygavhsplqapgryfehyrsiininrrryaamlscldeainnvtlalk 289
                                            264 AYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLALK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Example 4; Page 557-558; 894pp; English.
                                                                                                                                                    AAU14185 standard; Protein; 366 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Fang YT, Liu C, Drmanac RT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25-JAN-2001; 2001WO-US02623.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25-JAN-2000; 2000US-0491404
                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                             Human novel protein #56.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2001-451939/48.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (HYSE-) HYSEQ INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-PSDB; AAS22490
                                                                                                                                                                                                                                                                                                                                                                                                                                WO200155437-A2.
                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens.
                                                                                                                                                                                                               24-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                              02-AUG-2001.
                                                                                                                                                                                  AAU14185;
                                                                                                                       RESULT
AAU14185
               a
                                                                         g
                                                                                                                                                                                    ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Human, PRO; EST; expressed sequence tag; PCR primer; hybridisation; probe; blood coagulation disorder; cancer; cellular adhesion disorder; secreted protein; transmembrane protein.
                                                                                                                                                                                                                                      282 ihtglqhsiirptqpnclpldnatlpqklkevgysthmvgkwhlgfyrkecmptrrgfdt 341
                                                                                                                                                   PHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITGKYQ 135
                                                                                                                                                                     136 IHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDT 195
                                                                                                   QACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLAQAGEKLEPSTTSTSQ 75
                                              Gaps
                                                .;
0
            Length 366;
                                              Indels
                              Pred. No. 1.3e-190;
              Score 196; DB 22;
35.6%; Scor.
100.0%; Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                              AAY41701 standard; Protein; 515 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human PRO708 protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98US-0078936.
98US-0078939.
98US-0079294.
98US-0079656.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98US-0079663.
98US-0079664.
98US-0079689.
98US-0079728.
98US-0079920.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98US-0080105.
98US-0080107.
98US-0080165.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98US-0077641.
98US-0077649.
98US-0077791.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98US-0078886.
98US-0078910.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                98US-0080194
98US-0080327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99WO-US05028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98US-0078004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                        196 FFGSLLGSGDYYTHYK 211
                                                                                                                                                                                                                                                                                                            Best Local Similarity 100.
Matches 196; Conservative
```

```
polypeptides
                                                                                       for generation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QAGEKLEPSTISTSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQP 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Human; secreted protein; transmembrane protein; PRO; EST; cytostatic; expressed sequence tag; detection; cancer.
                                              The present invention describes secreted and transmembrane polypeptid and their polynucleotides. The nucleotide sequences are useful as sources of probes, primers, for chromosome mapping, and for generatio of antisense sequences. They can also be used to create transgenic animals. The proteins can be used to treat a variety of diseases and disorders, depending on their function. Diseases that may be treated include blood coagulation disorders, cancers and cellular adhesion disorders. They may also be used to raise antibodies. AAZ33891 to AAZ34338, and AAY41685 to AAY4174 represent polynucleotide and polypeptide sequence given in the exemplification of the present
                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                               1 MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121 ICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLG
                                                                                                                                                                                                                                                                                                                                                                         ;
                                                                                                                                                                                                                                                                                                                                    Length 515;
                                                                                                                                                                                                                                                                                                                                                                       0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Human PRO708 (UNQ372) protein sequence SEQ ID NO:114.
                                                                                                                                                                                                                                                                                                                                  Score 181; DB 20; I
Pred. No. 3.4e-175;
                                                                                                                                                                                                                                                                                                                                                                         Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB44257 standard; Protein; 515 AA.
                                                                                                                                                                                                                                                                                                                   32.9%; Scc.
100.0%; Pre/
0; /
               Claim 12; Fig 43; 530pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99Wo-US05028
99US-0126753
99US-0130232
99US-0131445-
99US-0141037
99US-0141037
99US-0145698
99Wo-US28313
99Wo-US28551
99Wo-US28551
99Wo-US28555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2000WO-US04341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                            515 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200053756-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-APR-1999;
14-MAY-1999;
23-JUN-1999;
26-JUL-1999;
29-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           02-DEC-1999;
02-DEC-1999;
16-DEC-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            08-FEB-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             29-MAR-1999
21-APR-1999
                                                                                                                                                                                                                                                                                                                                                                       Matches 181;
                                                                                                                                                                                                                                           invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 F 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       f 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB44257;
                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB44257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Сp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               dd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New secreted and transmembrane polypeptides and their polynucleotides, useful for treating blood coagulation disorders, cancers and cellular adhesion disorders -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ٠,
ب
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Chen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Baker KP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Yuan J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gurney A,
                                                                                                                                                                                                                                                                        98US - 0082704
98US - 0082804
98US - 0082767
98US - 0083336
98US - 0083322
98US - 0083322
98US - 0083495
98US - 0083495
98US - 00835496
98US - 00835496
98US - 0083545
98US - 008354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0083559
98US-0084342
98US-0084414
98US-0084441
98US-0084641
98US-0084600
98US-0084627
98US-0084627
98US-0084637
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         980S - 0084649
980S - 0084643
980S - 0085338
980S - 0085339
980S - 0085573
980S - 0085580
980S - 0085580
980S - 0085580
980S - 0085689
980S - 0085689
980S - 0085697
                                                                                   98US-0081195.
98US-0081203.
98US-0081229.
                                                                                                                                                                                                                 98US-0082568.
98US-0082569.
98US-0082700.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98US-0086023.
98US-0086392.
98US-0086414.
98US-0086430.
                             98US-0081049.
                                                                                                                                                               98US-0081838
98US-0081952
                                                                  98US-0081071
                                                                                                                                                                                                  98US-0081955
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98US-0086486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98US-0087098
                                                                                                                                           98US-0081817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98US-0087208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98US-0100038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GETH ) GENENTECH INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Goddard A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1999-551358/46.
N-PSDB; AAZ33987.
                                                                                                                                                                                                                                                                                                                                                                 28 - APR - 1998

29 - APR - 1998

30 - APR - 1998

31 - APR - 1998

32 - APR - 1998

33 - APR - 1998

34 - APR - 1998

35 - APR - 1998

36 - APR - 1998

37 - APR - 1998

38 - APR - 1998

39 - APR - 1998

39 - APR - 1998

30 - APR - 1998

31 - A
         01-APR-1998;
08-APR-1998;
08-APR-1998;
08-APR-1998;
09-APR-1998;
09-APR-1998;
                                                                                                                                        15-APR-1998;
15-APR-1998;
15-APR-1998;
15-APR-1998;
21-APR-1998;
22-APR-1998;
                                                                                                                                                                                                                                                                        22-APR-1998;
22-APR-1998;
23-APR-1998;
23-APR-1998;
27-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18-MAY-1998;
22-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               22-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11-SEP-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Wood WI,
```

Ŋ

us-09-495-823-7.oli.rag

```
Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pan J,
    ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRO polypeptide; manmal; tumour; cancer; human; cattle; horse; sheep; dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha; blood; chondrocyte cell; cell proliferation; cell differentiation; colon; adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
                                                                                                                                                                                                                                                                                sequence tag) sequences which encode secreted or transmembrane PRO polypeptides. The PRO polynucleotides and polypeptides have cytostatic activity. The polypeptides and polypeptides can be used for detecting the presence of PRO polypeptides in samples, for linking bioactive molecules to cells and for modulating biological activities of cells, using the polypeptides for specific targeting. The polypeptide targeting and be used to Aill the target cells, e.g. for the treatment of cancers. The polypeptide pairs provide specific targeting of bioactive molecules to cells. AAC78600 to AAC78987 represent PCR primers and probes used in the isolation of the PRO polynucleotide.sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Novel PRO polypeptides and polynucleotides used in detection methods, to target bloactive molecules to specific cells, and to modulate cellular activities -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QAGEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQP 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEGALLA 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
Kljavin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;
Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;
                                                                                                                                                                                                                                                                     AAC78458 to AAC78599 represent polynucleotide and EST (expressed
                                                                                                                                                                                                                                                                                                                                                                                                                                               32.9%; Score 181; DB 21; Length 515; 100.0%; Pred. No. 3.4e-175; Live 0; Mismatches 0; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Human PRO polypeptide sequence #38.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ŗ.
                                                                                                                                                                                                                                               Claim 12; Fig 43; 636pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAU29061 standard; Protein; 515
30-DEC-1999; 99WO-US31243.
30-DEC-1999; 99WO-US31274.
05-JAN-2000; 2000WO-US00219.
06-JAN-2000; 2000WO-US002776.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 32.9
Best Local Similarity 100.
Matches 181; Conservative
                                                                      (GETH ) GENENTECH INC
                                                                                                                                                                WPI; 2000-611443/58.
N-PSDB; AAC78483.
                                                                                                                                                                                                                                                                                                                                                                                                              515 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    18-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       F 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 f 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAU29061;
                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ő
```

```
Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention. The PRO polypeptides and their associated nucleic acids can be used to detect the presence of a tumour in a mammal by comparing the level of expression of a PRO polypeptide in a test sample of cells from the animal and a control sample of normal cells, whereby a higher level of expression in the test sample indicates the presence of a tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats and rabbits but are preferably human. The polypeptides can be used to stimulate tumour necrosis factor (TNF) alpha release from human blood, when contacted with it. A specific polypeptide can be used to stimulate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Novel nucleic acids encoding PRO polypeptides, used to diagnose the presence of tumours, such as prostate and breast tumours, in mammals and to screen for modulators of the compounds -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Godowski PJ, Gurney AL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Desnoyers L, Goddard A, Go
tanabe CK, Wood WI, Zhang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 11; Fig 76; 774pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ', Chen J, Desnoyers L
Smith V, Watanabe CK,
                                                                                                                                                                                                                                                            2000Wo-USO6884.
2000US-199104P.
2000US-191048P.
2000US-191314P.
2000US-191314P.
2000US-193032P.
                                                                                                                                                                                                                                                                                                                                                                                                                                  2000W3-US08439
2000US-194449P
2000US-194647P
2000US-195975P
2000US-196000P
2000US-196000P
2000US-196690P
2000US-196690P
                                                                                                                                                                        2000US-187202P.
2000US-186968P.
2000US-189320P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2000WO-US23328.
2000WO-US30952.
2000WO-US32678.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2000US-199654P
2000US-201516P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2000WO-US15264
2000US-209832P
                                                                                                                                                                                                                                          2000US-189328P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000US-198121P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2000US-199397P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2000US-199550P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000WO-US13705
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2000WO-US14042
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2000WO-US14941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2000US-0644848
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2000WO-US34956
                                                                                    28-FEB-2001; 2001WO-US06520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GETH ) GENENTECH INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2001-602746/68.
N-PSDB; AAS45962.
WO200168848-A2.
                                                                                                                                                                                          06-MAR-2000; 2
14-MAR-2000; 2
14-MAR-2000; 2
21-MAR-2000; 2
21-MAR-2000; 2
21-MAR-2000; 2
21-MAR-2000; 2
21-MAR-2000; 2
                                                                                                                                                                                                                                                                                                                                                                                         29-MAR-2000; 2
29-MAR-2000; 2
04-AAR-2000; 2
04-ARR-2000; 2
11-APR-2000; 2
11-APR-2000; 2
11-ARR-2000; 2
11-ARR-2000; 2
11-ARR-2000; 2
11-ARR-2000; 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      18-APR-2000; 2
25-APR-2000; 2
25-APR-2000; 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-APR-2000;
03-MAY-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 17-MAY-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          30-MAY-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                02-JUN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     05-JUN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JUL-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     22-MAY-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08-NOV-2000;
                                           20-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Baker KP,
```

```
Isolated human polynucleotides encoding polypeptides, useful for treatment and diagnosis of e.g. cancer, ulcers and HIV infection
                 N-PSDB; AAH99356
      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
Antificiammatory; antirheumatic; antiarthritic; immunosuppressive;
Antibacterial; endocrine; cardiant; central nervous system; virucide;
antibacterial; endocrine; cardiant; central nervous system; virucide;
anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemic;
antiaggregant; haemostatic; vulnerary; antilocr; osteopathic; eczema;
dermatological; antiallergic; antiathmatic; cytostatic;
neuroprotective; antidepressant; nootropic; antidiabetic; cytostatic;
neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
mutianaphylactic; rheumatolid arthritis; septic shock; pancreatitis;
cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
denetic disease; haematopoletic disorder; platelet disorder; asthma;
thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
allergic rhinitis; diabetes; mutitiple sclerosis; depression;
Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
                                                                                                                                                                                 ö
the proliferation or differentiation of chondrocyte cells. The PRO proteins can be used to determine the presence of tumours and also susceptibility to tumour development, particularly adrenal, lung, colon, breast, prostate, rectal, cervical, or liver tumours, in mammalian subjects. The oligonucleotide probes specific for the PRO nucleic acids can be used for genetic analysis of individuals with genetic disorders.
                                                                                                                                                                                                                                                                                       QAGEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKTPTLDKLAAEGVKLENYYVQP 120
                                                                                                                                                                                                                                                                                                                                 121 ICTPSRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLG 180
                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                            1 MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEGALLA
                                                                                                                                                                                 ö
                                                                                                                                                    Length 515;
                                                                                                                                                                                 Indels
                                                                                                                                                  32.9%; Score 181; DB 22; L 100.0%; Pred. No. 3.4e-175; tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Human protein sequence SEQ ID NO:930.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAM25415 standard; Protein; 187 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Liu C, Drmanac RT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23-DEC-1999; 99US-0471275.
21-JAN-2000; 2000US-0488725.
25-APR-2000; 2000US-0552317.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     22-DEC-2000; 2000WO-US35017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  neurological disorder.
                                                                                                                                                                  Similarity
                                                                                                        515 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (HYSE-) HYSEQ INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16-0CT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26-JUL-2001
                                                                                                                                                                  Best Local Sim
Matches 181;
                                                                                                                                                                                                                                                                                                                                                                                            181 F 181
                                                                                                                                                                                                                                                                                                                                                                                                                       181 É 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAM25415;
                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Tang YT,
                                                                                                                                                     Ouery Match
                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAM25415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                       q
   8888888888
                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                              à
```

```
CC AAH99166 to AAH99904 encode the human proteins given in AAM25255 to CAAM25963. The proteins can have activities based on the tissues and cells they are expressed in, such activities based on the tissues and cells they are expressed in, such activities based on the tissues and cells they are expressed in, such activities antinflammatory; antirheumatic; contrain nervous system; virucide; anti-HV; fungicide; antimutagen; cardiovascular; antianaemic; antiadegregant; haemostatic; vulnerary; antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic; antiparkinsonian; and immunostimulant. The proteins and polynucleotides cording them can be used in gene therapy, antisense therapy and vaccine conding them can be used in gene therapy, antisense therapy and vaccine condition; The proteins and polynucleotides are useful for screening for aponists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatold arthritis, septic shock, pancreatitis, cardiac dysfunction, rheumatold arthritis, severe combined immunodeficiency, eczenn inflammation, autoimmunity, genetic diseases, haematopoietic disorders, anematoporosis, severe combined immunodeficiency, eczenn allergic chinitis, asthma, diabetes, cancer, multiple sclerosis, depression, contrological disorders, neurodegenerative and contrological disease, neurodegenerative and contrological disease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase; ligase; hyperproliferative disorder; immunodeficiency disorder; autoimmune disorder; neurological disorder; metabolic disorder; inflammatory disorder; cardiovascular disorder; reproductive disorder; blood-related disorder; infectious disorder; cytostatic; anti arthritic; nephrotropic; anticoagulant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               134 YQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGF 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9 ygihtglqhsiirptqpnclpldnatlpqklkevgysthmvgkwhlgfyrkecmptrrgf 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            254 NPTKPIFLYIAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLDEAINNVTLALK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 187;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        32.5%; Score 179; DB 22; L
100.0%; Pred. No. 1.5e-173;
ive 0; Mismatches 0;
Claim 20; Page 200; 1217pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Novel human enzyme polypeptide #421.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAU23335 standard; Protein; 128 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 32.5
Best Local Similarity 100.
Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   187 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200155301-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              17-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   02-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAU23335;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAU23335
```

PF	17-JAN-2001;	2001WO-US01239.
XX PR	-JAN-	00003-(
PR	-FEB-	)-S0000
P.R.	24 - FEB - 2000;	2000US-0184664. 2000US-0186350
P.R.	-MAR-	)-SD000
۳. c	-MAR-	)-sn000
7 Z	-MAY	)-SD000
PR	- NDC -	)-sn000
7 G	NOD	00003-0
PR	-JUL-	)-S0000
PR :	-30L-	)-S0000
7 Z	- 100 -	)-SD000
PR:	105	)-S0000
PR c	,-105-	)-Sn000
7 G	-700F-	)-S0000
PR	-Aug-	)-snooo
PR G	-Aug-	)-Sn000
2. G	- AUG -	)-SD000
P. P.	-AUG-:	)-snooo
PR	-AUG-	)-Sn000
۳. و و	-AUG-	)-SD000
. A	-AUG-	-50000
PR	-AUG-	)-S0000
PR	-AUG-	0000s-(
PR G	-AUG-	)-S0000
7 0 2 0	- AUG-	- 50000
P. P.	- AUG -	- S0000
PR	-AUG-	0000s-(
PR	-Aug-	-S0000
P. 0	-SEP-	)-S11000
7 G	SEP-	-50000
P.R.	-SEP-	0000s-(
PR I	-SEP-	-S0000
۳ a	-SEP-	-SD000
PR	-SEP-	- S0000
PR	-SEP-	000us-(
PR c	-SEP-	)-S0000
	SEP	- 51000
PR	-SEP-	0000s-(
F 1	-SEP-	)-S0000
7. U	SEP	-50000
PR	-SEP-	0000s-(
PR :	-SEP-	)-SD000
7 G	SEP	)-S11000
PR:	-SEP-	)-SD000
PR	-SEP-	-sn000
۳. د د	-SEP-	0000s-(
ች ጜ	SEP-	)-S0000
PR	-SEP-	- SD000
PR :	-SEP-	)-S0000
7 G	SEP	- SILOOO
PR	-SEP-	0000s-(
PR o	-SEP-	)-SD000
PR	-SEP-	)-S0000
PR a	-SEP-	)-S0000
۲,	SEF	- 50000

PR 29-5EP-2000; 2000US-02286802.
PR 02-0CT-2000; 2000US-0237038.
PR 02-0CT-2000; 2000US-0237038.
PR 02-0CT-2000; 2000US-0237038.
PR 13-0CT-2000; 2000US-0237039.
PR 13-0CT-2000; 2000US-0239935.
PR 13-0CT-2000; 2000US-0241803.
PR 20-0CT-2000; 2000US-0241803.
PR 20-0CT-2000; 2000US-0241809.
PR 20-0CT-2000; 2000US-0246524.
PR 20-0CT-2000; 2000US-024929.
PR 20-0CT-2000; 2000US-0224929.
PR

```
The present invention relates to the isolation of novel human enzyme polypeptides, and the CDNA (AAS40785-AAS41684) and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the disorders including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g. athhills), meurological disorders (e.g. Alzhamatory disorders (e.g. athhills), neurological disorders (e.g. atheroselers), ce.g. athhills), neurological disorders (e.g. atheroselers), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atheroselerosis), blood related disorders (e.g. hemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. Influenza). The polynucleotides of the invention can also be used in gene therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                      the printed
                 preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous
                                                                                                                                                                                                                                                                                                                                                                                  e: The sequence data for this patent did not form part of the printer scification, but was obtained in electronic format directly from WIPO ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sulfatase; 25278 sulfatase; nootropic; neuroprotective; antibacterial; antiinflammatory; vasotropic; antitumor; gene therapy; human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     90 DVGYHGSEIKTPTLDKLAAEGVKLENYYVQPICTPSRSQFITGKYQIHTGLQHSIIRPTQ 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
0
and polynucleotides useful for diagnosing,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 128;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFYRKECMPTRRGFDTFFG 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Tsai F;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 109; DB 22; I
Pred. No. 1.8e-102;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Rudolph-Owen LA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mismatches
                                                                                   Claim 11; SEQ ID No 1331; 1180pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19.8%; Sco.
100.0%; Pred
0; N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Human 25278 sulfatase polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB85482 standard; Protein; 569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Glucksmann MA, Williamson M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (MILL-) MILLENNIUM PHARM INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            31-JAN-2001; 2001WO-US03266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          31-JAN-2000; 2000US-0495823.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2001-476214/51.
N-PSDB; AAH46862, AAH46866.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
polypeptides
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Α̈́
                                                                                                                                                                                                                                                                                                                                                                                                                                                      128
                                                                                                                                                                                                                                                                                                                                                                                      Note: The seque
specification,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO200155411-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          02-AUG-2001,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 109;
                                                                                                                                                                                                                                                                                                                                                                        invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB85482;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local S
 Novel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
 à
```

```
ö
                                                                                             The invention provides 22438, 23533, 25278 or 26212 human sulfatase polypeptides and polynucleotides. The sulfatase genes and polypeptides are useful for treating disorders involving the brain such as cerebrovascular diseases, infections such as acute meningitis, demyelinating diseases including multiple sclerosis, degenerative diseases affecting the cerebral cortex including Alaheimer's disease and Pick diseases, spinocerebellar degenerations including spinocerebellar ataxias including Friedreich ataxia, and ataxia telangiectasia, degenerative diseases affecting motor neurons including amyotrophic lateral sclerosis, inborn errors of metabolism such as leukodystrophies, toxic and acquired metabolic diseases, including vitamin deficiencies, and neurocutaneous syndromes (phakamatoses) including neurofibromotosis. The present sequence represents a human 25278 sulfatase polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The present invention describes a recombinant human iduronate 2-sulfatase
Novel human sulfatase polypeptides useful for treating and diagnosing sulfatase-related disorders such as cerebrovascular diseases, acute meningitis, multiple sclerosis, degenerative diseases and tumor
                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New highly glycosylated recombinant human iduronate 2-sulfatase useful for diagnosing or treating subjects suspected of having o suffering from IDS deficiency disorders, e.g. Hunter syndrome
                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Clements PR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Human; iduronate 2-sulfatase; IDS; glycosylated; gene therapy; mucopolysaccharidosis inhibitor; IDS deficiency disorder; Hunter syndrome; mucopolysaccharidosis type II.
                                                                                                                                                                                                                                                                                                                                                                                    Length 569;
                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bielicki J,
                                                                                                                                                                                                                                                                                                                                                                                  3.5%; Score 19; DB 22; I
100.0%; Pred. No. 2.7e-10;
Live 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Human sulfatase protein B SEQ ID NO:13.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (WOME-) WOMEN'S & CHILDREN'S HOSPITAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Occhiodoro T,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB51184 standard; Protein; 533 AA
                                                                   Claim 9; Fig 10; 180pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Example 1; Fig 3; 53pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92US-0991973.
94US-0345212.
91US-0790362.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      168 GYSTHMVGKWHLGFYRKEC 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0249003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (mucopolysaccharidosis-II)
                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity 100.
nes 19; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Morris CP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Anson DS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2001-060076/07
                                                                                                                                                                                                                                                                                                                                       569 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12-FEB-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hopwood JJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           17-DEC-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-NOV-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21-MAR-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US6153188-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28-NOV-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Wilson PJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB51184;
                                                                                                                                                                                                                                                                                                                                       Seguence
                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10
                                                                                                                                                                                                                                                                                                                                                                                                     Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB51184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of
(IDS). The recombinant IDS is more highly glycosylated than the naturally
             occurring enzyme isolated from human 11ssue. The recombinant human TDS can be produced in Chinese Hamster Ovary (CHO) cells or in a human cell. The recombinant IDS comprises a fusion protein. It is a mucopolysaccharidosis inhibitor and can be used in gene therapy. The recombinant IDS is useful in treating and diagnosing subjects suffering from or suspected of having IDS deficiency disorders, e.g. Hunter syndrome (mucopolysaccharidosis type II). The present sequence represents a human sulfatase which is used in comparison with the IDS sequence in an example from the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell
                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; SEQ ID NO 25161; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                     1.8%; Score 10; DB 22; Length 533; 100.0%; Pred. No. 0.37; ive 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Drosophila melanogaster polypeptide SEQ ID NO 25161.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Li PWD, Myers EW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABB66123 standard; Protein; 217 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               23-MAR-2001; 2001WO-US09231.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23-MAR-2000; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Drosophila melanogaster.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (ABB57737-ABB72072).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Adams M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2001-656860/75.
                                                                                                                                                                                                                                                                                                                            171 THMVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                               (PEKE ) PE CORP NY.
                                                                                                                                                                                                                                                                      Best Local Similarity
Matches 10; Conserv
                                                                                                                                                                                                  533 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-PSDB; ABL10226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO200171042-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          interactions -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             27-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Venter JC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABB66123;
                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                      ABB66123
                                                                                                                                                                                                                                                                                                                                                              g
5556666666668
                                                                                                                                                                                                                                                                                                                            ò
```

217 AA;

Sequence

```
The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella neumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to indentify proteins used in proliferation, to expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence represents an essential prokaryotic cellular proliferation protein.
                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Trawick JD, Carr GJ;
                                         Gaps
                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οţ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                    Pseudomonas aeruginosa cellular proliferation protein #311.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           New polynucleotides for the identification and development antibiotics, comprise sequences of antisense nucleic acids
         Length 217;
                                       Indels
                                                                                                                                                                                                                                                                                               Antisense; prokaryotic cellular proliferation protein; antibiotic; antibacterial; drug design.
                                       ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Haselbeck R, Ohlsen KL, Zyskind JW, Wall D,
Yamamoto RT, Xu HH;
         DB 22;
. 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          511pp; English.
                                       Mismatches
         Score 8;
Pred. No.
                                                                                                                                                                          AAU36321 standard; Protein; 330 AA.
1.5%; Scc.
100.0%; Pre
0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2000US-242578P.
2000US-253625P.
2000US-257931P.
2001US-269308P.
                                                                                                                                                                                                                                                                                                                                                                                                                                     21-MAR-2001; 2001WO-US09180.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2000US-191078P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000US-206848P.
2000US-207727P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Example 3; Seq ID No 11914;
                                                                                                                                                                                                                                      (first entry)
                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (ELIT-) ELITRA PHARM INC
                                                                                                                                                                                                                                                                                                                                             Pseudomonas aeruginosa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2001-611495/70.
         Query Match
Best Local Similarity
Matches 8; Conserv
                                                                     ERVDLSNR 511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAS54180
                                                                                                35 ervdlsnr 42
                                                                                                                                                                                                                                                                                                                                                                           WO200170955-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23-OCT-2000;
27-NOV-2000;
22-DEC-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21-MAR-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23-MAY-2000;
26-MAY-2000;
                                                                                                                                                                                                                                      14-FEB-2002
                                                                                                                                                                                                                                                                                                                                                                                                       27-SEP-2001
                                                                                                                                                                                                        AAU36321;
                                                                                                                                               12
                                                                     504
                                                                                                                                                             AAU36321
                                                                                                                                               RESULT
                                                                                                qq
                                                                                                                                                                                                          δ
```

Sequence

g

ABB63962;

13

셤 ö

```
TIDS). The recombinant IDS is more highly glycosylated than the naturally occurring enzyme isolated from human tissue. The recombinant human IDS occurring enzyme isolated from human tissue. The recombinant human IDS or an be produced in Chinese Hamster Ovary (CHO) cells or in a human IDS. The recombinant IDS comprises a fusion protein. It is a mucopolysaccharidosis inhibitor and can be used in gene therapy. The recombinant IDS is useful in treating and diagnosing subjects suffering from or suspected of having IDS deficiency disorders, e.g. Hunter syndrome (mucopolysaccharidosis type II). The present sequence represents a sea urchin sulfatase which is used in comparison with the IDS sequence in an example from the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                present invention describes a recombinant human iduronate 2-sulfatase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New highly glycosylated recombinant human iduronate 2-sulfatase (IDS) useful for diagnosing or treating subjects suspected of having or suffering from IDS deficiency disorders, e.g. Hunter syndrome
                                                                                                                                                                         Human; iduronate 2-sulfatase; IDS; glycosylated; gene therapy; mucopolysaccharidosis inhibitor; IDS deficiency disorder; Hunter syndrome; mucopolysaccharidosis type II.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 551;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Drosophila melanogaster polypeptide SEQ ID NO 16530.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Occhiodoro T, Bielicki J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 22;
                                                                                                                                          Sea urchin sulfatase protein U SEQ ID NO:15.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1.5%; Score 8; E
100.0%; Pred. No.
ive 0; Mismatch
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (WOME-) WOMEN'S & CHILDREN'S HOSPITAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ABB63246 standard; Protein; 579 AA
                                AAB51186 standard; Protein; 551 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Example 1; Fig 3; 53pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                92US-0991973.
94US-0345212.
91US-0790362.
                                                                                                                                                                                                                                                                                                                                                                99US-0249003.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (mucopolysaccharidosis-II)
                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Anson DS, (
Morris CP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2001-060076/07.
                                                                                                                                                                                                                                                       Arbacia punctulata.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         551 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     173 MVGKWHLG 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12-FEB-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       26-MAR-2002
                                                                                                      21-MAR-2001
                                                                                                                                                                                                                                                                                                                                                                                                    17-DEC-1992;
28-NOV-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                      12-NOV-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8;
                                                                                                                                                                                                                                                                                         US6153188-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hopwood JJ,
                                                                                                                                                                                                                                                                                                                             28-NOV-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ABB63246;
                                                                  AAB51186;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ABB63246
               AAB51186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DX LX XX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        to an isolated nucleic acid detection reagent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABLJ6176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins
                                                                                                                                                                                                                                                                                                                                                                                                      developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; SEQ ID NO 18678; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 542;
                                                   Length 330;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                  Drosophila melanogaster polypeptide SEQ ID NO 18678.
                                                                                       ;
0
                                                 1.5%; Score 8; DB 22;
100.0%; Pred. No. 26;
ive 0; Mismatches (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1.5%; Score 8; DB 22
100.0%; Pred. No. 41;
:1ve 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ξ×
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Myers
                                                                                                                                                                                                                                                         ABB63962 standard; Protein; 542 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PWD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23-MAR-2000; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23-MAR-2001; 2001WO-US09231
                                                                                                                                                                                                                                                                                                                               26-MAR-2002 (first entry)
                                  Query Match
Best Local Similarity luv..
E. . . . 8; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity 100.
nes 8; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ĽŢ
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Drosophila melanogaster.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        invention relates
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Adams M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2001-656860/75.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (ABB57737-ABB72072)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (PEKE ) PE CORP NY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           542 AA;
330 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         120 PICTPSRS 127
                                                                                                                                                N-PSDB; ABL08065
                                                                                                                           53 EEEGALLA 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO200171042-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                           pharmaceutical
                                                                                                                                                                                                                                                                                                                                                                                                        Drosophila;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JC,
```

Venter

Clements

ö

Gaps

ö

Indels

Sequence

Query Match

Best Loc Matches

7.1

ŏ g

```
The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sequences (ABLO1840-ABL16175) and the encoded proteins (ABB7737-ABB72072).
The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                  New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions - \,
               Drosophila; developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                                                                                                     Disclosure; SEQ ID NO 16530; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                    Li PWD, Myers EW;
                                                                                                                                                                                                23-MAR-2000; 2000US-191637P. 11-JUL-2000; 2000US-0614150.
                                                                                                                                                                23-MAR-2001; 2001WO-US09231
                                                             Drosophila melanogaster.
                                                                                                                                                                                                                                                                                  Venter JC, Adams M,
                                                                                                                                                                                                                                                                                                                WPI; 2001-656860/75.
N-PSDB; ABL07349.
                                                                                                                                                                                                                                                  (PEKE ) PE CORP NY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      579 AA;
                               pharmaceutical.
                                                                                              WO200171042-A2
                                                                                                                               27-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
```

Search completed: July 23, 2002, 19:08:41 Job time: 347 sec

ö

Gaps

. 0

Length 579; 0; Indels

DB 22;

Query Match 1.5%; Score 8; DB 2 Best Local Similarity 100.0%; Pred. No. 43; Matches 8; Conservative 0; Mismatches

174 VGKWHLGF 181 

ò

	. *		
			•

colon cancer PRO708 hybri PRO708 hybri

encoding

polynucleoti cancer assoc

colon cancer immune/haema

immune/haema

Perfect score: Sequence: Scoring table:

OM nucleic

Run on:

Word size :

Database

Searched:

```
CDNA encoding nove
Human PRO708 PCR f
Human PRO708 PCR r
Human PRO708 forwa
Human PRO708 rever
                                                                                                                                                                           Human polynucleoti
Human secreted pro
Apoptosis related
Human immunoglobul
Human secreted pro
Human cDNA SEQ ID
Human cDNA SEQ ID
Human immunoglobul
                                                                                                                                                                                                                                                                        Human immune/haema
DNA encoding human
                                                                                                                                                                                                                                                                                                 Human ADAM protein
              Human colon cancer
                                                                                                                                                                                                                                                                                                             P. chrysogenum ABC
                                                                                                                                                                                                                                                                                                                       Human genomic DNA
Human genomic DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sulfatase; 23553 sulfatase; nootropic; neuroprotective; antibacterial; antiinflammatory; vasotropic; antitumor; gene therapy; human; ss.
                                                                                                                                                       Human
                                                                                                                                                                                                                                                                                                                                                                                                       Human
                                                                                                                                                                                                                                                                                                                                                                                 Human
                                                                       Human
                                                                                   Human
                           Human
                                                            Human
                                                                                                                                                                                                                                                                                                                                               Human
                                                                                                                                                                                                                                                                                                                                                            Human
                                                                                                                                                                                                                                                                                                                                                                       Human
                                                                                                                                                                                                                                                                                                                                                                                              Human
                                                                                                                                                                                                                                                                                                                                                                                                                      Human
                                      Human
                                                 Human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Human 23553 sulfatase polypeptide encoding cDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       "23553 sulfatase"
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ALIGNMENTS
                                                                                                                                                                                                              AAS28772
AAD05310
ABA06520
                                                                                                                                                                                                                                                                       AAK61788
AAS33072
                                                                                                                                                                                                                                                                                              AAD18169
AAF83398
AAS26779
                                                                                                                             AAC78673
AAC78674
AAH36471
                                                                                                                                                                                                                                                                                                                                                                              AA190981
AA192024
AA193263
                                                                                                       AAZ33988
AAZ33989
                                                                                                                                                                AAK57023
AAI83302
                                                                                                                                                                                        AAX00620
AAA95796
                                                                                                                                                                                                                                                                                                                                  AAS26780
AAK81080
                                                                                                                                                                                                                                                                                                                                                         AAK56582
ABA08227
                                                                                                                                                                                                                                                              AAS28838
                                   AAZ33990
                                                                                AAC77614
                                                                                                                                                                                                                                                  ABA06709
                                                                                                                                                                                                                                                                                                                                                                                                                   AA185187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers 324.,2123
/*tag= a /product= "23553 sul:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAH46863 standard; cDNA; 2253 BP.
 31-JAN-2001; 2001WO-US03266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     31-JAN-2000; 2000US-0495823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
--000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200155411-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              25-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      02-AUG-2001
AAH46863;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAH46863
  (without alignments)
10907.615 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Human 2353 sulfat
Human 2353 sulfat
Human DNA encoding
Human PRO708 (UNQ3
Human PRO708 nucle
Human cDNA encodin
Human cDNA encodin
                                                                                                                                                                                                                                                                                                                                      N_Geneseq_032802:*

| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Description
                                                                                                                                                    1 cacgogtccgcccacgcgtc.....tgccacctggtgccgaattc
                                                                             July 23, 2002, 15:52:14 ; Search time 356.68 Seconds
                                                                                                                                                                                                                                                           3472872
           4.5
Compugen Ltd.
                                                                                                                                                                                                                                                        Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                             1736436 segs, 858457221 residues
         GenCore version
Copyright (c) 1993 - 2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                                                                                                                                                                                                                                                    Post-processing: Listing first 45 summaries
                                                       - nucleic search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAH46863
AAH46867
AAS45962
AAC78483
AAC323726
AAS22490
AAH99356
AAA02344
                                                                                                                                                                          OLIGO_NUC
Gapop 60.0 , Gapext 60.0
                                                                                                                                                                                                                                                                                Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                            US-09-495-823-8
2266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2253
1800
4639
4640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4650
1375
1340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              82.8
68.4
667.3
67.3
67.3
67.3
26.0
23.7
21.1
                                                                                                                                                                                                                                  0
```

polynucleoti polynucleoti polynucleoti

Tsai F;

Rudolph-Owen LA,

Glucksmann MA, Williamson M, (MILL-) MILLENNIUM PHARM INC

protein enco colon cancer

Human Human

1877 1524 1526 1526 1526 890 590 537 479

Score

Result Š.

immune/haema CDR62 homolo

polynucleoti

immune/haema

99

2

501

acute

and

treating

The invention provides 22438, 23553, 25278 or 26212 human sulfatase polypeptides and polynucleotides. The sulfatase genes and polypeptides are useful for treating disorders involving the brain such as cerebrovascular diseases, infections such as acute meningitis, demyelinating diseases including multiple sclerosis, degenerative diseases including multiple sclerosis, degenerative diseases sffecting the cerebral cortex including Alzheimer's disease and Pick disease, spinocerebellar degenerations including spinocerebellar ataxias including Friedreich ataxia, and ataxia telangiectasia, degenerative diseases affecting motor neurons including amyotrophic lateral sclerosis, inborn errors of metabolism such as leukodystrophies, toxic and acquired metabolic diseases, including vitamin deficiencies, and neurocutanneous syndromes (phakamatoses) including neurofibromotosis. sulfatase-related disorders such as cerebrovascular diseases, a meningitis, multiple sclerosis, degenerative diseases and tumor useful for Novel numan sulfatase polypeptides sulfatase-related disorders such as Claim 2; Fig 15; 180pp; English 2001-476214/51 P-PSDB; AAB85483 encoding 

T; 0 other; 540 . છ 554 ; ; Sequence 2253 BP; 624 A; 535

1; 300 300 360 360 420 420 480 540 900 099 agccctccttgttcttccggagtcccatccattaagccatcacttctggaagattaaagt 180 Gaps 9 9 ggaagctgctctagggaggggggggaggaggagaaagtgaaaatgtgctggagaagagcg agteteacetgggaeetgaatgaatggeteceaggggetgtgegggggaetecgeeteege accacggatctgagattaaaacacctactcttgacaagctcgctgccgaaggagttaaac tgtcggacatggtgacagctgagaggaggaggatttcttgccaggtggagagtcttca cogtetgttgggtgcatgtgtgcgccgcascggcgcgggggcgcgtggttetecgcgtgg agteteacetgggacetgaatgaatggeteceaggggetgtgegggggateeteege tctggatcctctgcctcacttatggttacctgtcctgggggccaggccttagaagagg aggaagaaggggccttactagctcaagctggagagagaactagagcccagcacaacttcca aggaaggagggccttactagctcaagctggagagaaactagagcccagcacaacttcca cotoccagococcatotoattttcatcotagoggatgatcagggatttagagatgtgggtt ; Length 2253; 5; Indels 22; DB Mismatches Score 1877; Pred. No. 0; ; 0 82.8%; 99.7%; Query Match 82.8 Best Local Similarity 99.7 Matches 2247; Conservative Н Н 61 121 181 241 61 121 181 241 301 541 301 361 361 121 121 181 481 541 601 g õ Ω Q ò g ò õ g ò ò g à g δ a ò g ò Op ò

960 780 780 840 900 900 accacggatctgagattaaaacacctactcttgacaagctcgctgccgaaggagttaaac acaatggcatatactccacacagatgtacactcagagagtacagcaaatcttagcttccc aagctcctggcaggtatttcgaacactaccgatccattatcaacataaacaggaggagat aageteetggeaggtatttegaacaetacegateeattateaacataaacaggaggagat atgaggacattcaactagatggctatgatatctgggagaccataagtgagggtcttcgct atgaggacattcaactagatggctatgatatctgggagaccataagtgagggtcttcgct caccccgagtagatattttgcataacattgaccccatatacaccaaggcaaaaatggct caccccgagtagatattttgcataacattgaccccatatacaccaaggcaaaaatggct cctgggcagcagcatctggatctggaacactgcaatccagtcagccatcagagtgcagc tacctctggacaatgccaccctacctcagaaactgaaggaggttggatattcaacgcata acagtcctgggatgtgtggctatgacttgtatgaaaacgacaatgctgcctgggactatg atgetgecatgetteetgettagatgaageaateaacaacgtgaeattggetetaaaga atgetgecatgettteetgettagatgaageaateaacaacgtgacattggetetaaaga cttatggtttctataacaacagcattatcatttactcttcagataatggtggccagccta cttatggtttctataacaacagcattatcatttactcttcagataatggtggccagccta cggcaggagggagtaactggcctctcagaggtagcaaaggaacatattgggaaggagga aacctgtgcacatcactgactggtaccccactctcatttcactggctgaaggacagattg cctgggcagcaggctatgggatctggaacactgcaatccagtcagccatcagagtgcagc actggaaattgcttacaggaaatcctggctacagcgactgggtcccccctcagtctttca tacctctggacaatgccaccctacctcagaaactgaaggaggttggatattcaacgcata 1141 1081 1081 1201 1381 1501 1681 1141 1201 1261 1261 1321 1381 1441 1021 1021 1321 1501 1561 1561 1621 1681 199 199 721 721 781 781 841 841 901 901 961 1621 g g g qq qq qq op οy g δ g δy g δλ Db δ g δ Б Qγ g δ qq ò Qγ g ŏ á g ö ò à q ò ò ŏ

 $\sim$ 

Wed

```
gacttgtatgaaaacgacaatgctgcctgggactatgacaatggcatatactccacacag 1043
                                                                                                                                  Similarity
                                                                                                   coding sequence.
                                                                                                                             Query Match
                                                                                                                                   Local
                                                                                                                                                   324
                                                                                                                                                                                                                            181
                                                                                                                                                                        384
                                                                                                                                                                                                                                                 241
                                                                                                                                                                                                                                                                                                                                                           541
                                                                                                                                                                                                                                                                                                                                                                                           984
                                                                                                                                                                                            444
                                                                                                                                                                                                       121
                                                                                                                                                                                                                                       564
                                                                                                                                                                                                                                                                      301
                                                                                                                                                                                                                                                                                           361
                                                                                                                                                                                                                                                                                                                 421
                                                                                                                                                                                                                                                                                                                                                                      924
                                                                                                                                                                                  61
                                                                                                                                                                                                                  504
                                                                                                                                                                                                                                                            624
                                                                                                                                                                                                                                                                                 684
                                                                                                                                                                                                                                                                                                      744
                                                                                                                                                                                                                                                                                                                            804
                                                                                                                                                                                                                                                                                                                                      481
                                                                                                                                                                                                                                                                                                                                                864
                                                                                                                                                                                                                                                                                                                                                                                601
                                                                                                                                                                                                                                                                      Dp
                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                g
P
                                                                                                                                                                        ōλ
                                                                                                                                                                                  셤
                                                                                                                                                                                            ò
                                                                                                                                                                                                       d
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                           ď
                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                            ð
                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                         2040
                                                                                                             2099
                                                                                                                        2100
                                                                                                                                  2159
                                                                                                                                                        2219
   1980
                                                                        2160
                                                   Sulfatase; 23553 sulfatase; nootropic; neuroprotective; antibacterial; antiinflammatory; vasotropic; antitumor; gene therapy; human; ss.
                                                                                                            caggaatcgtgaagaagctcctacggaggctctcacagttcaacaaaactgcagtgccgg
                                                                                         catggtatagagagaaaccaagaaaaagaagccaagcaaaaatcaggctgagaaaaagc
                                                                                              tcaggtatcccccaaaagaccccagaagtaaccctaggctcaatggaggggtctagggac
                                                                                                                                                                                                                                                                                                                                                                      ٠.
ن
                                                                                                                                                                                                                                                                                                                                                                      Tsai
                                                                                                                                                                                                                                                Human 23553 sulfatase polypeptide coding sequence.
                                                                                                                                                                                                                                                                                                                                                                      ľĄ,
                                                                                                                                                                                                                                                                                                                                                                      Rudolph-Owen
                                                                                                                                                                           /*tag= a
/product= "23553 sulfatase"
                                                                                                                                                                                                                                                                                    Location/Qualifiers
1..1800
                                                                                                                                                                                                                 BP
                                                                                                                                                                                                                                                                                                                                                                      Glucksmann MA, Williamson M,
                                                                                                                                                                                                                 AAH46867 standard; cDNA; 1800
                                                                                                                                                                                                                                                                                                                                                           (MILL-) MILLENNIUM PHARM INC
                                                                                                                                                                                                                                                                                                                                     31-JAN-2001; 2001WO-US03266
                                                                                                                                                                                                                                                                                                                                                31-JAN-2000; 2000US-0495823
                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                               WPI; 2001-476214/51.
P-PSDB; AAB85483.
                                                                                                                                                                                                                                                                                                                WO200155411-A2
                                                                                                                                                                                                                                                                            sapiens
                                                                                                                                                                                                                                      25-SEP-2001
                                                                                                                                                                                                                                                                                                                           02-AUG-2001
                                                                                                                                                                                                                           AAH46867;
              1741
                                                                                                            2041
                                                                                                                       2041
                                                                                                                                                       2160
                                                                                                                                                                 2161
                                                                                                                                                                            2220
                                                                                                                                                                                       2221
                                                                                                                                                                                                                                                                           Ношо
                                   1801
                                                                   1921
                                                                             1921
                                                                                                  1981
                                                                                                                                  2100
                                                                                                                                            2101
                        1801
                                              1861
                                                        1861
                                                                                       1981
                                                                                                                                                                                                                                                                                      Key
                                                                                                                                                                                                            AAH46867
    ö
              a
                                  g
                                                        g
                                                                             셤
                                                                                       ò
                                                                                                g
                                                                                                            õ
                                                                                                                      g
                                                                                                                                           g
                                                                                                                                                       ò
                                                                                                                                                                 a
                                                                                                                                                                            ò
                                                                                                                                                                                      g
                                                                                                                                                                                                                       ò
                                              õ
                                                                   ò
                                                                                                                                  ò
```

```
cerebrovascular diseases, infections such as acute meningitis, enquyelinating diseases including multiple sclerosis, degenerative diseases including multiple sclerosis, degenerative diseases affecting the cerebral cortex including Alzheimer's disease and Pick disease, spinocerebellar degenerations including spinocerebellar ataxias including Friedracion ataxia, and ataxia telangiectusia, degenerative diseases affecting motor neurons including amyotrophic lateral sclerosis, inborn errors of metabolism such as leukodystrophic vice and acquired metabolic diseases, including vitemin deficiencies, and neurocutaneous syndromes (phakamatoses) including neurofibromotosis. The present sequence represents a human 13553 sulfatase polypeptide
                                                                                                                                   sulfatase
polypeptides
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 743
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 863
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    900
Novel human sulfatase polypeptides useful for treating and diagnosing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 atggctcccaggggctgtgcgggggatccgcctccgccttctccacaggcctgtgtctgt 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       923
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                983
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                       such as cerebrovascular diseases, acute
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          cctggaaagatgctagcaatgggggcgctggcaggattctggatcctctgcctccact
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        atcctagcggatgatcagggatttagagatgtgggttaccacggatctgagattaaaaca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      cctcagaaactgaaggaggttggatattcaacgcatatggtcggaaaatggcacttgggt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ttttacagaaaagaatgcatgcccaccagaagaggatttgataccttttttggttccctt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ttgggaagtggggattactatacacactacaaatgtgacagtcctgggatgtgtggctat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        tatggttacctgtcctggggccaggccttagaagaggaggaagaaggggccttactagct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             atcctagcggatgatcagggatttagagatgtgggttaccacggatctgagattaaaaca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 atttgcacaccatccaggagtcagtttattactggaaagtatcagatacacaccggactt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       caacattctatcataagacctacccaacccaactgtttacctctggacaatgccacccta
                                          meningitis, multiple sclerosis, degenerative diseases and tumor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 1800;
                                                                                                                            The invention provides 22438, 23553, 25278 or 26212 human s polypeptides and polynucleotides. The sulfatase genes and pare useful for treating disorders involving the brain such perphynosemula.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 1800 BP; 534 A; 434 C; 423 G; 409 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 1549;
Pred. No. 0;
                                                                                  Claim 2; Fig 15; 180pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        68.48;
99.98;
                     sulfatase-related disorders
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 1649; Conservative
```

ö

encoding PRO polypeptide sequence #38

(first entry)

4

```
18-DEC-2001
        Human DNA
 1343
                                                        1020
                                                              1403
                                                                   1080
                                                                         1463
                                                                              1140
                                                                                   1260
                                                                                                         1643
                                                                                                                   1703
                                                                                                                         1380
                                                                                                                                   1440
                                                                                                                                         1823
                   1163
                              1223
                                        1283
                                                                                                              1320
                                                                                                                              1763
                                                                                                                                              1500
                                                                                                                                                    1883
                                                                                                                                                                    1620
                                              960
                        840
                                   900
                                                             cacaatgaacggatcacctcgtcaactggcaaaagtgtatggcttttcaacatcacagcc
                                        gatgaagcaatcaacaacgtgacattggctctaaagacttatggtttctataacaacagc
                                                   attatcatttactcttcagataatggtggccagcctacggcaggaggagtaactggcct
                                                                                                                                              cacaatgaacggatcaccttgtcaactggcaaaagtgtatggcttttcaacatcacagcc
                                                                                                                                                                            agaagtaaccctaggctcaatggaggggtct 1974
                                                             1344
                                                                             1081
                                                   1284
                                                                   1021
                                                                                        1141
                                                                                              1524
                                                                                                   1201
                                                                                                        1584
                                                                                                              1261
                                                                                                                   1644
                                                                                                                                   1381
                                                                                                                                        1764
                                                                                                                                                                   1561
                                                                                                                                                                         1944
                                                                                                                                                                              1621
  661
             721
                  1104
                        781
                             1164
                                   841
                                        1224
                                             901
                                                        196
                                                                        1404
                                                                                   1464
                                                                                                                        1321
                                                                                                                              1704
                                                                                                                                              1441
                                                                                                                                                   1824
                                                                                                                                                         1501
                                                                                                                                                              1884
                                                                                                                                                                                       RESULT
AAS45962
        ò
             g
                  ö
                       g
                             ò
                                  g
                                      δ
                                            엄
                                                  ò
                                                       Ω
                                                            9
0
0
                                                                       ò
                                                                            QQ
                                                                                   oy
op
                                                                                            ŏ
                                                                                                  qq
                                                                                                        δý
                                                                                                            g
                                                                                                                  ò
                                                                                                                       g
                                                                                                                             ô
                                                                                                                                 g
                                                                                                                                        ò
                                                                                                                                              g
                                                                                                                                                        g
                                                                                                                                                                   g
                                                                                                                                                                              8
                                                                                                                                                   ò
                                                                                                                                                              ò
                                                                                                                                                                         ò
```

BP

AAS45962 standard; cDNA; 4639

Ωī

AAS45962

```
PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep; ss; dog; cat; pig; goat; rabblt; tumour necrosis factor alpha; TNF-alpha; blood; chondrocyte cell; cell proliferation; cell differentiation; colon; adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sed to diagnose the tumours, in mammals
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gurney
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ЪЈ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Godowsk1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  nucleic acids encoding PRO polypeptides, used noe of tumours, such as prostate and breast tum reen for modulators of the compounds -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ζ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A, Go
Zhang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Goddard Wood WI,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŗ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 2; Fig 75; 774pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Desnoyers |
tanabe CK,
                                                                                                                                                                                                                                                                                                                                                                                                               2000Wo-USO5601.
2000US-187202P
2000US-186968P
2000US-189320P
2000US-189328P
2000Wo-USO6884
2000US-190828P
2000US-191007P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000US-191314P
2000US-192655P
2000US-193032P
2000US-193053P
2000US-193053P
2000US-194449P
2000US-195975P
2000US-195975P
2000US-196690P
2000US-196690P
2000US-196820P
2000US-198121P
2000US-198131P
2000US-198585P
2000US-199337P
2000US-199337P
2000US-199554P
2000US-199554P
2000US-199554P
2000US-199554P
2000US-199554P
2000US-199554P
2000US-199554P
2000US-199554P
2000US-199554P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2000WO-US15264
2000US-209832P
2000US-0644848
2000WO-US20710
2000WO-US3952
2000WO-US3952
2000WO-US3956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Watanabe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GETH ) GENENTECH INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2001-602746/68.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Chen J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Smith V,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2001-602746
P-PSDB; AAU29061
                                                                                                                                                                                                                                      WO200168848-A2.
                                                                                                                                                                                                                                                                                                                                                                                                               01-MAR-2000; 203-MAR-2000; 204-MAR-2000; 206-MAR-2000; 206-MAR-2000; 214-MAR-2000; 215-MAR-2000; 215-MAR-2000; 221-MAR-2000; 221-MAR-2000; 221-MAR-2000; 221-MAR-2000; 221-MAR-2000; 221-APR-2000; 221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         02-JUN-2000;
05-JUN-2000;
28-JUL-2000;
22-AUG-2000;
                                                                                                                                                                                  Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24-AUG-2000;
08-NOV-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    20-DEC-2000;
                                                                                                                                                                                                                                                                                                                                                           28-FEB-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    presence of
                                                                                                                                                                                                                                                                                                 20-SEP-2001
                                                                                                                    PCR primer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  to screen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ΚP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pan J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Baker
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Novel
```

Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR primers for PRO polypeptides of the invention. The sequences of the invention can be used to detect the presence of a tumour in a mammal by comparing the level of expression of a PRO polypeptide in a test sample of cells from the animal and a control sample of normal cells, whereby a higher level of expression in the test sample indicates the presence of a tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep, bigs, goats and rabbits but are preferably human. The polypeptides can be used to stimulate tumour necrosis factor (TNF) alpha release from human blood, when contacted with it. A specific polypeptide can be used to stimulate the proliferation of differentiation of chondrocyte cells. The PRO proteins can be used to determine the presence of tumours and also susceptibility to tumour development, particularly adrenal, lung, colon, breast, prostate, rectal, cervical, or liver tumours, in mammalian changed the used for genetic analysis of individuals with genetic disorders. 9999999999999999988

Sequence 4639 BP; 1426 A; 956 C; 1025 G; 1232 T; 0 other;

1; 1188 ctatttgcacaccatccaggagtcagtttattactggaaagtatcagatacacaccggac 1308 261 321 888 501 ctcaagctggagagactagagcccagcacaacttccacctcccagccccatctcattt 561 741 Gaps gcgccccgcascggcgcgggggggcgtggttctccgcgtggagtctcacctgggacctgagt gegecegeagegegegegegegegegetetecegegegegeteteacetgggacetgagt gaatggctcccaggggctgtgcggggcatccgcctccgccttctccacaggcctgtct ctcaagctggagagaaactagagcccagcacaacttccacctcccagccccatct ctatttgcacaccatccaggagtcagtttattactggaaagtatcagatacacaccggac ttcaacattctatcataagacctacccaacccaactgtttacctctggacaatgccaccc gaatggctcccaggggctgtgcggggcatccgcctccgccttctccacaggcctgtgtct cttatggttacctgtcctggggccaggccttagaagaggaggaagaaggggccttactag 4639 Length Indels 22; 9 8 Score 1526; DB Pred. No. 0; 0; Mismatches 0; 67.3%; Conservative Similarity Best Local Simi Matches 1946; Query Match 1069 1189 22 589 82 649 202 92 829 949 1129 742 142 709 262 322 889 382 442 600 502 562 g QQ g g 셤 음 g g q ò ð g ð ò à ò g ô ò ò ò ò ò ò g

1368 1488 1548 1041 1101 1161 1341 1401 1461 1521 1581 1641 2207 1701 2267 1761 1881 921 981 atgacttgtatgaaaacgacaatgctgcctgggactatgacaatggcatatactccacac agatgtacactcagagagtacagcaaatcttagcttcccataaccccacaaagcctatat ccgacccatatgagaggtggacctatctaacaggtatccaggaatcgtgaagaagctcc atgacttgtatgaaaacgacaatgctgcctgggactatgacaatggcatatactccacac agatgtacactcagagagtacagcaaatcttagcttcccataaccccacaaagcctatat gcattatcatttactcttcagataatggtggccagcctacggcaggaggagtaactggc gctatgatatctgggagaccataagtgaggtcttcgctcaccccgagtagatattttgc gctatgatatctgggagaccataagtgagggtcttcgctcaccccgagtagatattttgc atcctggctacagcgactgggtccccctcagtctttcagcaacctgggaccgaaccggt gtttttacagaaaagaatgcatgcccaccagaagaggatttgatacctttttggttccc ttttatatattgcctatcaagctgttcattcaccactgcaagctcctggcaggtatttcg tacctcagaaactgaaggaggttggatattcaacgcatatggtcggaaaatggcacttgg 2208 1762 1822 1309 1369 1429 1489 1549 1609 1102 1669 1162 1729 1222 1789 1282 1849 1342 1909 1402 1969 1462 2029 1522 2089 1582 2149 1642 1702 2328 802 862 922 982 1042 Dp qq g g g δλ g οy qq δ q QY g OD Óγ g ŏ QQ ò g οy a δ ٥y a οý ò ŏ δ q ò ò οy

```
Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                  646
                                                                                                                                                                                                                                                                                                                                                                                                                         502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1429
                                                                                                                                                                                                                                                                                                                              889
                                                                                                                                                                                                                                                                                                                                                                                                                                          1069
                                                                                                                                                                                                    142
                                                                                                                                                                                                                                                                                                                                                                                    442
                                                                                                                                                                                                                                                                                                                                                                                                                                                            562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1369
                                                                                                                                                                                 549
                                                                                                                                                                                                                     607
                                                                                                                                                                                                                                                        694
                                                                                                                                                                                                                                                                                           829
                                                                                                                                                                                                                                                                                                                                                382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      742
                                                                                                                            22
                                                                                                                                                                82
                                                                                                                                                                                                                                       202
                                                                                                                                                                                                                                                                          262
                                                                                                                                                                                                                                                                                                             322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Dp
88888888888
                                                                                                                                              g
                                                                                                                                                                                  a
                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                             염
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                 δ
                                                                                                                                                                                                    à
                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAC78458 to AAC78599 represent polynucleotide and EST (expressed sequence tag) sequences which encode secreted or transmembrane PRO polypeptides. The PRO polynucleotides and polypeptides have cytostatic activity. The polynucleotides and polypeptides can be used for detecting the presence of PRO polypeptides in samples, for linking bloactive
cegaeccetatgagaggtggaectatetaacaggtatecaggaategtgaagaagetee 2447
                                                                                                                                                                                          Human; secreted protein; transmembrane protein; PRO; EST; cytostatic;
expressed sequence tag; detection; cancer; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Novel PRO polypeptides and polynucleotides used in detection methods, to target bloactive molecules to specific cells, and to modulate cellular activities -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Æ;
                 tacggaggctctcacagttcaacaaaactgcagtgccggtcaggtatcccccaaagacc
                                    tacggaggctctcacagttcaacaaaactgcagtgccggtcaggtatccccccaaagacc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gerber H, Gerritsen I
rney AL, Hillan KJ;
Paoni NF, Roy MA;
ms PM, Wood WI;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Desnoyers L, Eaton DL;
                                                                                                                                                                         Human PRO708 (UNQ372) nucleotide sequence SEQ ID NO:113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gurney AL,
J, Paoni NF,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Napier MA, Pan J, Paoni N
TA, Tumas D, Williams PM,
                                                   Botstein D, Der
Fong S, Gao W,
Grimaldi CJ, G
                                                                                                                     ВÞ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 2; Fig 42; 636pp; English.
                                                                                                                     CDNA; 4640
                                                                                                                                                                                                                                                                                                   99US-0123957.
99US-0126773.
99US-0131445.
99US-0134287.
99US-0141037.
99US-0145698.
99US-0162508.
                                                                                                                                                                                                                                                                                                                                                                                    99WO-US28551.
99WO-US28565.
99WO-US30095.
                                                                                                                                                                                                                                                                                                                                                                                                             99WO-US31243.
                                                                                                                                                                                                                                                                         2000WO-US04341
                                                                                                                                                                                                                                                                                                                                                                                                                                2000WO-US00219
                                                                                                                                                                                                                                                                                                                                                                                                                                                  2000WO-US00376
                                                                                                                                                                                                                                                                                                                                                                                                                                         2000WO-US00277
                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ashkenazi AJ, Baker KP,
Ferrara N, Filvaroff E,
Goddard A, Godowski PJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Stewart TA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GETH ) GENENTECH INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Kuo SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2000-611443/58.
P-PSDB; AAB44257.
                                                                                                                   AAC78483 standard;
                                                                                                                                                                                                                                      WO200053756-A2
                                                                                                                                                                                                                                                                                                                                                                 29-0CT-1999;
30-0NOV-1999;
02-DEC-1999;
16-DEC-1999;
30-DEC-1999;
30-DEC-1999;
05-JAN-2000;
06-JAN-2000;
06-JAN-2000;
                                                                                                                                                                                                                      Homo sapiens
                                                                                                                                                                                                                                                                         18-FEB-2000;
                                                                                                                                                                                                                                                                                          08-MAR-1999;
12-MAR-1999;
29-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Kljavin IJ,
Shelton DL,
                                                                                                                                                                                                                                                                                                                                               23-JUN-1999,
26-JUL-1999
                                                                                                                                                       08-FEB-2001
                                                                                                                                                                                                                                                        14-SEP-2000
                                                                                                                                                                                                                                                                                                                       21-APR-1999
                                                                                                                                                                                                                                                                                                                              28-APR-1999
                                                                                                                                                                                                                                                                                                                                        14-MAY-1999
                                                                                                                                     AAC78483;
                                                                                                  4
                                  2448
                                                    1942
                                                                      2508
2388
                  1882
                                                                                                         AAC78483
                                                                                                                            QQ
                                                                      g
                 ð
                                                     ö
```

```
1;
molecules to cells and for modulating biological activities of cells, using the polypeptides for specific targeting. The polypeptide targeting can be used to kill the target cells, e.g. for the treatment of cancers. The polypeptide pairs provide specific targeting of bloactive molecules to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1068
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      741
                                                                                                                                                                                                                                                                                          648
                                                                                                                                                                                                                                                                                                                               141
                                                                                                                                                                                                                                                                                                                                                               708
                                                                                                                                                                                                                                                                                                                                                                                                  201
                                                                                                                                                                                                                                                                                                                                                                                                                                    768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             801
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                861
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       921
                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                      81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gcgcccgcascggcgggggggggggtggttctccgcgtggagtctcacctgggacctgagt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gegecegeagegegegegegegtggttetecgegtggagteteacetgggacetgagt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gaatggctcccaggggctgtgcgggcatccgcctccgccttctccacaggcctgtgtct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       gaatggctcccaggggctgtgcggggcatccgcctccgccttctccacaggcctgtgtct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1009 cttatggttacctgtcctggggccaggccttagaagaggaagaagaaggggccttactag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ctcaagctggagagaaactagagcccagcacaacttccacctcccagccccatctcattt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ttcaacattctatcataagacctacccaacccaactgtttacctctggacaatgccaccc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ttcaacattctatcataagacctacccaacccaactgtttacctctggacaatgccaccc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                tacctcagaaactgaaggaggttggatattcaacgcatatggtcggaaaatggcacttgg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   tacctcagaaactgaaggaggttggatattcaacgcatatggtcggaaaatggcacttgg
                                                                                                                                                                                                                                                                                                                               ggaggaggaggagaaagtgaaatgtgctggagaagagggggggccctccttgttcttccgga
                                                                                                                                                                                                                                                                                                                                                                                                    gtcccatccattaagccatcacttctggaagattaaagttgtcggacatggtgacagctg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               cttatggttacctgtcctggggccaggccttagaagaggaggaagaaggggccttactag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           tcatcctagcggatgatcagggatttagagatgtgggttaccacggatctgagattaaaa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       gtttttacagaaaagaatgcatgcccaccagaagaggatttgatacctttttggttccc
                                                                                                                                                                                                                                                        gtcctggaaagatgctagcaatggggggggctggcaggattctggatcctctgcctcctca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1,
                                                                                                                             Sequence 4640 BP; 1427 A; 955 C; 1026 G; 1232 T; 0 other;
                                                                                                                                                                                   Length
                                                                                                                                                                                                                      Indels
                                                                                         the isolation of the PRO polynucleotide sequences.
                                                                                                                                                                                 21;
                                                                                                                                                                                                                   9
                                                                                                                                                                                   DВ
                                                                                                                                                                                                                    0; Mismatches
                                                                                                                                                                                 Score 1526;
Pred. No. 0;
                                                                                                                                                                                 67.3%;
99.6%;
                                                                                                                                                                                                                      Matches 1946; Conservative
                                                                                                                                                                                                    Similarity
```

BP.

CDNA; 4650

standard;

/

```
AAZ33987
RESULT
AAZ33987
         1461
                                                                                                                                      1641
1788
                                                                   1848
                                                                         1401
                                                                                                                    2088
                                                                                                                                2148
                                                                                                                                            2207
                                                                                                                                                        2267
                                                                                                                                                                                 2387
                                     1161
                                        aacactaccgatccattatcaacataaacaggagagatatgctgccatgcttcctgct 1221
                                                             1281
                                                                                                                          1581
                                                                                                                                                  1701
                                                                                                                                                               1761
                                                                                                                                                                           1821
                                                                                                                                                                                                          2507
                                                                                     aacactaccgatccattatcaacataaacaggagaagaatatgctgccatgctttcctgct
                                                             tagatgaagcaatcaacaacgtgacattggctctaaagacttatggtttctataacaaca
                                                                                                             gotatgatatotgggagaccataagtgagggtcttcgctcaccccgagtagatattttgc
                                                                                                                                      ataacattgaccccatatacaccaaggcaaaaaatggctcctgggcagcaggctatggga
                                                                                                                                         ttttatatattgcctatcaagctgttcattcaccactgcaagctcctggcaggtatttcg
                                                                                                                                                                                                               1489
                  1549
                                                                   1789
                                                                                                                         1522
                                                                                                                                                                          1762
                                                                                                                                                                                                                     2508
                        1042
                              1609
                                    1102
                                          1669
                                                1162
                                                      1729
                                                             1222
                                                                         1282
                                                                               1849
                                                                                     1342
                                                                                           1909
                                                                                                  1402
                                                                                                        1969
                                                                                                              1462
                                                                                                                   2029
                                                                                                                               2089
                                                                                                                                      1582
                                                                                                                                            2149
                                                                                                                                                  1642
                                                                                                                                                        2208
                                                                                                                                                               1702
                                                                                                                                                                    2268
                                                                                                                                                                                 2328
                                                                                                                                                                                             2388
                                                                                                                                                                                                   1882
                                                                                                                                                                                                         2448
                                                                                                                                                                                                               1942
922
            982
                                                                                                                                                                                       1822
     셤
                 Q
                             g
                                         g
                                                                                           g
                                                                                                      q
                                                                                                             ò
                                                                                                                   셤
                                                                                                                               q
                                                                                                                                           Q
                                                                                                                                                       g
                                                                                                                                                                   g
                                                                                                                                                                                a
                                                                                                                                                                                            a
                                                                                                                                                                                                         d
à
           ò
                        ò
                                   õ
                                                ò
                                                     q
                                                            ò
                                                                  g
                                                                         ò
                                                                              셤
                                                                                     ογ
                                                                                                 õ
                                                                                                                        õ
                                                                                                                                     οy
                                                                                                                                                 ŏ
                                                                                                                                                               ŏ
                                                                                                                                                                          δ
                                                                                                                                                                                       ò
                                                                                                                                                                                                   ð
                                                                                                                                                                                                               ò
```

```
Human, PRO; EST; expressed sequence tag; PCR primer; hybridisation; . probe; blood coagulation disorder; cancer; cellular adhesion disorder; secreted protein; transmembrane protein; ss.
                                                                nucleotide sequence.
                                                                                                                                                                                                                                                                                                                                                  9805-0077791
9805-0078004-
9805-0078086-
9805-0078936-
9805-0078936-
9805-0078936-
9805-0079664-
9805-0079664-
9805-0079664-
9805-0079728-
9805-0079728-
9805-0079728-
9805-0079728-
9805-0079728-
9805-0079733-
9805-0080105-
9805-0080105-
9805-0080107-
9805-0080107-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
9805-00801070-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98US-0083392.
98US-0083495.
98US-0083496.
                                                                                                                                                                                                                                                                                     -0077450.
-0077632.
-0077641.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98US-0083336
                                                                                                                                                                                                                                                         99WO-US05028
                               (first entry)
                                                                                                                                                                                                                                                                                     980S-
980S-
980S-
980S-
                                                                PR0708
                                                                                                                                                             sapiens
                                                                                                                                                                                          WO9946281-A2
                                                                                                                                                                                                                                                                                   110 - MAR - 1998;

111 - MAR - 1998;

112 - MAR - 1998;

113 - MAR - 1998;

20 - MAR - 1998;

21 - MAR - 1998;

22 - MAR - 1998;

23 - MAR - 1998;

24 - MAR - 1998;

27 - MAR - 1998;

27 - MAR - 1998;

27 - MAR - 1998;

31 - MAR - 1998;

32 - APR - 1998;

33 - APR - 1998;

34 - APR - 1998;

35 - APR - 1998;

37 - APR - 1998;

38 - APR - 1998;

39 - APR - 1998;

31 - APR - 1998;

31 - APR - 1998;

32 - APR - 1998;

33 - APR - 1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         29-APR-1998;
29-APR-1998;
29-APR-1998;
                               07-DEC-1999
                                                                                                                                                                                                                                                      08-MAR-1999
                                                                                                                                                                                                                          16-SEP-1999
AAZ33987
                                                                Human
                                                                                                                                                             Ношо
```

us-09-495-823-8.oli.rng

648

ggaggaggagaaagtgaaatgtgctggagaagaggcccccccttgttcttccgga

589 82 201 768 261

549

qq

δ

607

qq

142

ŏ

691

임

202

ŏ

829 322 889

g ŏ qq ŏ g ŏ

ŏ

agaggagagaggatttcttgccaggtggagagtcttcaccgtctggttgggtgcatgtg gcgcccgcascggcgcgggggcgcgtggttctccgcgtggagtctcacctgggacctgagt

gaatggctcccaggggctgtgcggggcatccgcctccgccttctccacaggcctgtgtct

321

œ

```
The present invention describes secreted and transmembrane polypeptides and their polynucleotides. The nucleotide sequences are useful as sources of probes, primers, for chromosome mapping, and for generation of antisense sequences. They can also be used to create transgenic animals. The proteins can be used to treat a variety of diseases and disorders, depending on their function. Diseases that may be treated include blood coagulation disorders, cancers and cellular adhesion disorders. They may also be used to raise antibodies. AAX33891 to AAX34338, and AAY41774 represent polynucleotide and polypeptide sequence given in the exemplification of the present
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   secreted and transmembrane polypeptides and their polynucleotides, ful for treating blood coagulation disorders, cancers and cellular
                                                                                                                                                                                                                                                                                                                                                                                                                                                ņ
                                                                                                                                                                                                                                                                                                                                                                                                                                                Chen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0 other
                                                                                                                                                                                                                                                                                                                                                                                                                                                ΚP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ..
H
                                                                                                                                                                                                                                                                                                                                                                                                                                                Baker
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C; 1026 G; 1235
                                                                                                                                                                                                                                                                                                                                                                                                                                                ٦,
                                                                                                                                                                                                                                                                                                                                                                                                                                                Yuan
                                                                                                                                                                                                                                                                                                                                                                                                                                                Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 2; Fig 42; 530pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BP; 1431 A; 958
                                                                                                                                                                                                                                                                                                                                                                                                                                                Gurney
                                                                                                                                                                                                                           98US-0085579.
98US-0085580.
98US-0085582.
                                                                                                                                                                                                                                                                                                                                                          98US-0087098.
                                                                                                                                                                                                                                                                                                                                                                                98US-0087208.
           98US-0083500
                                                   98US-0083559
98US-0083742
                                                                         98US-0084366
                                                                                                        98US-0084598
                                                                                                                     98US-0084600
                                                                                                                              98US-0084627
98US-0084637
                                                                                                                                                   98US-0084639
                                                                                                                                                                        98US-0084643
                                                                                                                                                                                                      98US-0085339
98US-0085573
                                                                                                                                                                                                                                                           98US-0085689
                                                                                                                                                                                                                                                                                98US-0085700
                                                                                                                                                                                                                                                                                                                                     98US-0086430
                                                                                                                                                                                                                                                                                                                                                98US-0086486
                                                                                                                                                                                                                                                                                                                                                                                                     98US-0100038
                                                                                              98US-008441
                                                                                                                                                                                                                                                                      98US-0085697
                                                                                                                                                                                                                                                                                           98US-0085704
                                                                                                                                                                                                                                                                                                     98US-0086023
                                                                                                                                                                                                                                                                                                                          98US-0086414
                                                                                                                                                                                 98US-0085323
                                                                                                                                                                                                                                                                                                                                                                                                                          (GETH ) GENENTECH INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                Goddard A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1999-551358/46.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          for treation disorders
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              P-PSDB; AAY41701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 4650
                   29-APR-1998;
29-APR-1998;
29-APR-1998;
                                                 29-APR-1998
30-APR-1998
66-MAY-1998
66-MAY-1998
07-MAY-1998
07-MAY-1998
07-MAY-1998
07-MAY-1998
07-MAY-1998
13-MAY-1998
13-MAY-1998
13-MAY-1998
13-MAY-1998
15-MAY-1998
15-MAY-1998
15-MAY-1998
15-MAY-1998
15-MAY-1998
15-MAY-1998
15-MAY-1998
                                                                                                                                                                                                                                                                                                    18-MAY-1998;
22-MAY-1998;
22-MAY-1998;
22-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                22-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                     11-SEP-1998
                                                                                                                                                                                                                                                                                                                                                                      28-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                28-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                               Wood WI,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        adheston
```

1008

441

1068

1009

ДQ

442

1069

g

502

οy

1129

622

295

Ω ద δ

ctcaagctggagagaactagagcccagcacaacttccacctcccagccccatctcattt ctcaagctggagagaaactagagcccagcacaacttccacctcccagccccatctcattt tcatcctagcggatgatcagggatttagagatgtgggttaccacggatctgagatttaaaa tcatcctagcggatgatcagggatttagagatgtgggttaccacggatctgagattaaaa cacctactcttgacaagctcgctgccgaaggagttaaactggagaactactatgtccagc

382 949

gaatggctcccaggggctgtgcggggcatccgcctccgccttctccacaggcctgtgtct gtcctggaaagatgctagcaatggggggggcttggcaggattctggatcctctgcctca 1188

1308

741

```
cacctactcttgacaagctcgctgccgaaggagttaaactggagaactactatgtccagc 1248
                                                           atgacttgtatgaaaacgacaatgctgcctgggactatgacaatggcatatactccacac
                                                                                                                                                                                                                                                atgacttgtatgaaaacgacaatgctgcctgggactatgacaatggcatatactccacac
                                                                                                                                                                                                                                                                  agatgtacactcagagagtacagcaaatcttagcttcccataaccccacaaagcctatat
                                                                                                                                                                                                                                                                                        agatgtacactcagagagtacagcaaatcttagcttcccataaccccacaaagcctatat
                      ctatttgcacaccatccaggagtcagtttattactggaaagtatcagatacacaccggac
                                                                                                     tacctcagaaactgaaggaggttggatattcaacgcatatggtcggaaaatggcacttgg
                                                                                                                                            gtttttacagaaaagaatgcatgcccaccagaagaggatttgatacctttttggttccc
                                                                                                                                                                                                                                                                                                           ttttatatattgcctatcaagctgttcattcaccactgcaagctcctggcaggtatttcg
 1189
                                                                                                    802
                                                                                                                                                              1429
                                                                                                                                                                                                        1489
                                                                                                                                                                                                                                              1549
                                                                                                                                                                                                                                                                                      1609
                                                                               1309
                                                                                                                       1369
                                                            742
                                                                                                                                           862
                                                                                                                                                                                  922
                                                                                                                                                                                                                            982
                                                                                                                                                                                                                                                                   1042
                     682
                                                                                                                                                                                                                                                                                                           1102
                                       g
                                                                                                                      g
                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                      g
g
                     ŏ
                                                            δ
                                                                              a
                                                                                                 ò
                                                                                                                                            ò
                                                                                                                                                              g
                                                                                                                                                                                  à
                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                  1;
                                                                                                                                                                                                                                                                                                 Gaps
```

1488

921

981

1161

1669

g

Length 4650;

Indels

9 DB 20;

0; Mismatches Score 1526; Pred. No. 0;

Matches 1946; Conservative

22

õ

Local Similarity

Query Match

67.3%; 99.6%;

σ

```
immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic; anticonvulsant; antiathritic; cerebroprotective; antidinugal; antiviral; antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic; thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
                                                                                                                             1461
                                                                                                                                                                                                                                        2088
                                                                                                                                                                                                                                                                                                                                                                                        1761
          1521
                                                                                                                                                                                                                                                                                                                           2207
                                                                                                                                                                                                                                                                                                                                              2387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1881
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       tacggaaggctctcacagttcaacaaaactgcagtgccggtcaggtatccccccaaagacc 1941
                                                                                                                                                                                                                                                             gctatgatatctgggagaccataagtgagggtcttcgctcaccccgagtagatattttgc 1581
                                                                                                                                                                                                                                                                                                       ataacattgaccccatatacaccaaggcaaaaatggctcctgggcagcagctatggga 1641
                                                                                                                                                                                                                                                                                                                                                                                                             atcetggetacagegaetgggtcecectcagtetttcageaacetgggacegaaceggt 2327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                protein; ss; Antianaemic; osteopathic; antiinflammatory;
                                        gcattatcatttactcttcagataatggtggccagcctacggcaggaggaggagtaactggc
                                                                                                                                                                                    atcctggctacagcgactgggtccccctcagtctttcagcaacctgggaccgaaccggt
                                                                                                                                                                                                                                                                                                                                                                                                                                             aacactaccgatccattatcaacataaacaggaggagatatgctgccatgctttcctgct
                                                                                                        gcattatcatttactcttcagataatggtggccagcctacggcaggaggagtaactggc
                                                                                                                                                                       atagcccacttctgaaaaacaagggaacagtgtgtaaggaacctgtgcacatcactgact
                                                                                                                                                                                                                 ggtaccccactctcatttcactggctgaaggacagattgatgaggacattcaactagatg
                                                                                                                                                                                                                                                                                                                ccgacccatatgagagggtggacctatctaacaggtatccaggaatcgtgaagaagctcc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ccgacccatatgagagggtggacctatctaacaggtatccaggaatcgtgaagaagctcc
                                                                                                                                                                                                                                                                                                                                                                                                                                   ggcacaatgaacggatcacctcgtcaactggcaaaagtgtatggcttttcaacatcacag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        encoding a novel human protein #292.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ВP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAS22726 standard; cDNA; 1375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                novel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Human cDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAS22726;
                                                                                                                                                 1909
                                                                                                                                                                                           1969
1162
                                         1222
                                                            1789
                                                                                                       1849
                                                                                                                                                                                                                                       2029
                                                                                                                                                                                                                                                          1522
                                                                                                                                                                                                                                                                                                     1582
                                                                                                                                                                                                                                                                                                                          2149
                                                                                                                                                                                                                                                                                                                                                                   2208
                                                                                    1282
                                                                                                                             1342
                                                                                                                                                                      1402
                                                                                                                                                                                                                 1462
                                                                                                                                                                                                                                                                                2089
                                                                                                                                                                                                                                                                                                                                               1642
                                                                                                                                                                                                                                                                                                                                                                                        1702
                                                                                                                                                                                                                                                                                                                                                                                                            2268
                                                                                                                                                                                                                                                                                                                                                                                                                                  1762
                                                                                                                                                                                                                                                                                                                                                                                                                                                      2328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1822
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2448
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAS22726
õ
                   g
                                                                                                                                               g
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                           q
                                                                                   ð
                                                                                                      q
                                                                                                                           ŏ
                                                                                                                                                                      ð
                                                                                                                                                                                         g
                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              쉽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
The invention relates to polynucleotides encoding novel human proteins or their active domains. The polypeptides, polynucleotides and antibodo parallel in a method of cartibodies raised against the polypeptides are used in a method of cartibodies raised against the polypeptides caused by the aberrant protein expression or activity. The polypeptides can be used as considered to identify compounds which bind to the polypeptides. Polynucleotides of the invention are used as probes and primers, for sequencing, for chromosome or gene mapping, in the production of recombinant proteins, and in generating anti-sense DNA or RNA and in gene therapy. Polypeptides of the invention can be used to carget drugs to a tumour, in assays to determine biological activity, to raise antibodies/elicit an immune response, to determine duantitative protein levels, as tissue markers, and to isolate receptors or ligands. Comparately the call disorders, regenerating bore, cartilage, tendon, cligament and/or nerve tissue, wound healing, treating platelet contraceptive, treating osteoporosis and osteoarthritis, anaemia, contraceptive, treating osteoporosis and osteoarthritis, anaemia, cungal infection or from autoinmunity, cancer, allergy, asthma, cantiland and infection. The manner of disease, and infection.

The profilammatory disease, eczema, heamophilia, thrombosis, and infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
 disorder; cancer; asthma; osteoporosis;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  428
                                                                                                                                                                                                                                                                                                                                                           Isolated polypeptides useful for treating anti-inflammatory diseases, nervous system disorders, and for regenerating bone and cartilage -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               anti-inflammatory diseases, nervous system disorders, and infection. The present sequence encodes a protein of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               caggcctgtgtctgtcctggaaagatgctagcaatggggggcgctggcaggattctggatc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  caggoctgtgtctgtcctggaaagatgctagcaatgggggggggcgctggcaggattctggatc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ctctgcctcctcacttatggttacctgtcctggggccaggccttagaagaggaagaa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ggggccttactagctcaagctggagagaaactagagcccagcacaacttccacctcccag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ggggccttactagctcaagctggagagaaactagagcccagcacaacttccacctcccag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 1375;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 1375 BP; 399 A; 336 C; 360 G; 280 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        39.3%; Score 890; DB
llarity 100.0%; Pred. No. 0;
Conservative 0; Mismatches
inflammatory di
immune disorder
                                                                                                                                                                                                                                                                                                                                                                                                               Claim 1; Page 699; 894pp; English.
                                                                                                                                                                                                                                                                     Liu C, Drmanac RT;
                                                                                                                                                           25-JAN-2001; 2001WO-US02623.
                                                                                                                                                                                              25-JAN-2000; 2000US-0491404
Parkinson's disease;
tissue regeneration;
                                                                                                                                                                                                                                                                                                     2001-451939/48.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
Matches 890; Conser
                                                                                                                                                                                                                                  HYSE-) HYSEQ INC
                                                                                                                                                                                                                                                                                                                         P-PSDB; AAU14421
                                                                                     WO200155437-A2
                                                      Homo sapiens.
                                                                                                                         02-AUG-2001
                                                                                                                                                                                                                                                                     Tang YT,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           489
   рp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δλ
```

RT;

Drmanac

Liu

YT,

Tang

HYSEQ INC ပဲ

(HYSE-)

2001-451939/48

P-PSDB; AAU14185

```
immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic; anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral; antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic; thrombolytic; immunogen; antibody; gene therapy; neurological disorder; Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis; tissue regeneration; immune disorder.
                                                                                                                                                                                                                                                                        1148
                                                                                                                                                                                                                                                                                                                                                                                                             ggcaggtatttcgaacactaccgatccattatcaacataaacaggaggagatatgctgcc 1208
                                                                                                                                                                                                                                         atatactccacacagatgtacactcagagagtacagcaaatcttagcttcccataacccc 1088
                                                                                                                                                                                                       964 aaatggcacttgggtttttacagaaaagaatgcatgcccaccagaagaggatttgatacc 1023
                                                                                                                                                                                   806
                                                                                                                                                                                                                             968
                                                                                                                903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            novel protein; ss; Antianaemic; osteopathic; antiinflammatory;
                                                                                       tctgagattaaaacacctactcttgacaagctcgctgccgaaggagttaaactggagaac
            gacaatgccaccctacctcagaaactgaaggaggttggatattcaacgcatatggtcgga
                                                                                                                                                            gacaatgccaccctacctcagaaactgaaggaggttggatattcaacgcatatggtcgga
                                                                                                                                                                                                                                                                                                                                                                 ggcaggtatttcgaacactaccgatccattatcaacataaacaggaggagatatgctgcc
                                            tactatgtccagcctatttgcacaccatccaggagtcagtttattactggaaagtatcag
                                                                                                                                                                                   aaatggcacttgggtttttacagaaaagaatgcatgcccaccagaagaggatttgatacc
                                                                                                                                                                                                                             ttttttggttcccttttgggaagtggggattactatacacactacaaatgtgacagtcct
                                                                                                                                                                                                                                                                                                                                                                                                                                                        cDNA encoding a novel human protein #56.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAS22490 standard; cDNA; 1340 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25-JAN-2001; 2001WO-US02623.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO200155437-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 02-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAS22490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Human;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Human
                                                                                                              844
                                                                                                                                                                                                                                                   1024
                                                                                                                                                                                                                                                                                                                     1029
                                                                                                                                                                                                                                                                                                                                                                1089
                                                                                                                                                                                                                                                                                                                                                                                                             1149
609
                     724
                                            699
                                                                  784
                                                                                         729
                                                                                                                                    789
                                                                                                                                                          904
                                                                                                                                                                                                                            606
                                                                                                                                                                                                                                                                         696
                                                                                                                                                                                                                                                                                                                                                                                                                                  1264
                                                                                                                                                                                849
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
AAS22490
                                                                                                                                                        g
                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                        ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                     qq
                                            ò
                                                               a
                                                                                         ö
                                                                                                             g
                                                                                                                                     ò
                                                                                                                                                                                ò
                                                                                                                                                                                                     g
                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                            q
```

25-JAN-2000; 2000US-0491404.

```
proteins or their active domains. The polypeptides, polynucleotides and national states to polynucleotides encoding novel human proteins or their active domains. The polypeptides, polynucleotides and antibodies raised against the polypeptides are used in a method of treatment of a mammal and prevention of disorders caused by the aberrant protein expression or activity. The polypeptides can be used as molecular weight markers, food supplements, and in antibody production. The polypeptides re used to identify compounds which bind to the polypeptides. Polynucleotides of the invention are used as probes and production of recombinant proteins, and in generating anti-sense DNA or RNA and in gene therapy. Polypeptides of the invention can be used to target drugs to a tumour, in assays to determine biological activity, to raise antibodies/elicit an immune response, to determine quantitative protein levels, as tissue markers, and to isolate receptors or ligands. Polypeptides of the invention may also be useful in treating platelet disorders, stem cell disorders, wound healing, treating burns, promoting the proliferation, differentiation and survival of stem cells, as contraceptive, treating osteoporosis and osteoarthritis, anaemia.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        723
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             668
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            783
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                          Isolated polypeptides useful for treating anti-inflammatory diseases,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            nervous system disorders, and infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ccccatctcattttcatcctagcggatgatcagggatttagagatgtgggttaccacgga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             tctgagattaaaacacctactcttgacaagctcgctgccgaaggagttaaactggagaac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              tctgagattaaaacacctactcttgacaagctcgctgccgaaggagttaaactggagaac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        atacacaccggacttcaacattctatcataagacctacccaacccaactgtttacctctg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        nervous system disorders, and for regenerating bone and cartilage
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ggggccttactagctcaagctggagagaaactagagcccagcacaacttccacctcccag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ccccatctcattttcatcctagcggatgatcaggggatttagagatgtgggttaccacgga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 tactatgtccagcctatttgcacaccatccaggagtcagtttattactggaaagtatcag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   tactatgtccagcctatttgcaccaccatccaggagtcagtttattactggaaagtatcag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       fungal infection or from autoimmunity, cancer, allergy, asthma, graft-versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory diseases, nervous system disorders, and infect
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 1340;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             he present sequence encodes a protein of the invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 1340 BP; 374 A; 338 C; 369 G; 258 T; 1 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26.0%; Score 590; DB 22; I 100.0%; Pred. No. 2.6e-272;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100.0%; Prea. ...
                                                                                                                                                                                              English.
                                                                                                                                                                                           Claim 1; Page 256-257; 894pp;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity
Matches 590; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      729
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
```

```
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19-MAY-2000
                                                                                                                                                                                                              Query Match
Best Local Si
Matches 537;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAA02344;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     σ
                                                                                                                                                                                                                                                                                                                                                                           842
                                                                                                                                                                                                                                                                                                                                                                                                                               902
                                                                                                                                                                                                                                                                                                                                                                                                                                                        204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1082
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1142
                                                                                                                                                                                                                                                                    722
                                                                                                                                                                                                                                                                                                                        782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   962
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
AAA02344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                               antibritammatory; antirheumatic; antiarthriti; immunosuppressive; antibraterial; endocrine; cardiant; central nervous system; virucide; antibraterial; endocrine; cardiant; central nervous system; virucide; antibraterial; hamostatic; vulnerary; antiulcer; osteopathic; eccema; dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic; antidiabetic; cytostatic; immunostimulant; gene therapy; antisense therapy; vaccine; inflammation; antianaphylactic; rheumatology cardiac anaphylaxis; autoimmunity; genetic disease; haematopoietic disorder; platelet disorder; asthma; thrombocytopaenia; osteoporosis; severe combined immunodeficiency; allergic rhinitis; diabetes; multiple sclerosis; depression;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAH99166 to AAH99904 encode the human proteins given in AAW25225 to AAW2963. The proteins can have activities based on the tissues and cells they are expressed in, such as: antinitammatory; antinheumatic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant; central nervous system; virucide; anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; antiagregant; heemostatic; vulnerary; antiulcer; osteopathic; dermatological; antiallergic; antiasthmatic; antidabetic; cytostatic; neuroprotective; antidepressant; nootropic; antiparkinsonian; and immunostimulant. The proteins and polynucleotides
                                                                                          aaatggcacttgggtttttacagaaaagaatgcatgcccaccagaagaggatttgatacc 1023
                                                                aaatggcacttgggtttttacagaaaagaatgcatgcccaccagaagaagtgttgatacc 908
                                     gacaatgccaccctacctcagaaactgaaggaggttggatattcaacgcatatggtcgga 963
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              the.
           gacaatgccaccctacctcagaaactgaaggaggttggatattcaacgcatatggtcgga
                                                                                                                                                                                                                                                                                                                     cancer; ulcer; HIV infection; human immunodeficiency virus;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Isolated human polynucleotides encoding polypeptides, useful for treatment and diagnosis of e.g. cancer, ulcers and HIV infection
                                                                                                                Human protein encoding cDNA sequence SEQ ID NO:191.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 1; Page 383; 1217pp; English.
                                                                                                                                                                                                              BP.
                                                                                                                                                                                                              AAH99356 standard; cDNA; 562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23-DEC-1999; 99US-0471275.
21-JAN-2000; 2000US-0488725.
25-APR-2000; 2000US-0552317.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22-DEC-2000; 2000WO-US35017
                                                                                                                                                                                                                                                                 16-OCT-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Fang YT, Liu C, Drmanac
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            neurological disorder; ss
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2001-457603/49.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (HYSE-) HYSEQ INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2001-457603/
P-PSDB; AAM25415.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200153455-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26-JUL-2001
                                                                                                                                                                                                                                       AAH99356;
                                                                                                                                                                                   œ
                                    904
                                                              849
                                                                                       964
                                                                                                                                                                                                 AAH99356
                                                                                       g
                                                                                                                                            g
                                                                                                                                                                                                                                       임
                                                                                                                  ò
            ò
                                                                ò
```

```
ö
encoding them can be used in gene therapy, antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmunity, genetic diseases, haematopoietic disorders, anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers, astemporosis, severe combined immunodeficiency, eczema, allergic rhintiis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 caatggcatatactccacacagatgtacactcagagagtacagcaaatcttagcttccca 1081
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        901
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          961
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           caatggcatatactccacacagatgtacactcagagagtacagcaaatcttagcttccca 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              841
                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Human colon cancer cell line polynucleotide sequence SEQ ID NO:2335.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Human; colon cancer; tumour; diagnosis; gene expression product; probe; detection; cancerous state; metastasis; identification; breast cancer; oestrogen receptor-positive breast cancer; therapy; oestrogen receptor-negative breast cancer; so.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        acctctggacaatgccaccctacctcagaaactgaaggaggttggatattcaacgcatat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            tgataccttttttggttcccttttgggaagtggggattactatacaccactacaatgtga
                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                           Length 562;
                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                   Sequence 562 BP; 170 A; 132 C; 105 G; 155 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                      23.7%; Score 537; DB 22; L
100.0%; Pred. No. 6.4e-247;
ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAA02344 standard; cDNA; 729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                         neurological disorders.
                                                                                                                                                                                                                                                                                                                                                                                                               al Similarity
537; Conserv
```

```
WO200155301-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         17-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             02-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAS41205;
                                                                                                                                                                                                                                                                                                                                                                                                10
                                                                                                                                   407
1745
                                                                                          1805
                                      Dp
                                                                                          ò
                                                                                                                                     g
                                                                                                                                                                                    ò
                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ibraries constructed from human colon cancer cell lines. The present libraries constructed from human colon cancer cell lines. The present invention also describes a method of detecting differentially expressed genes correlated with a cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample derived from a cell suspected of being cancerous, where detection of the differentially expressed gene product is correlated with a cancerous state of the cell from which the test sample was derived. The polynucleotides sequences can be used in a method for detecting differentially expressed genes correlated with a cancerous state of a mammalian cell. The polynucleotides can also be used as probes for detecting and mapping related genes. They can be used in diagnosis and promosis of diseases and disorders (e.g. identification of premetastatic or metastatic cancerous states, stages of cancer, or responsiveness of cancer to therapy). This is particularly for breast cancer, oestrogen receptor-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polynucleotide library used to determine cancerous states of mammalian
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1445 tgtgcacatcactgactggtaccccactctcatttcactggctgaaggacagattgatga 1504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ggacattcaactagatggctatgatatctgggagaccataagtgagggtcttcgctcacc 1564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ccgagtagatattttgcataacattgaccccatatacaccaaggcaaaaatggctcctg 1624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ggcagcaggctatgggatctggaacactgcaatccagtcagccatcagagtgcagcactg 1684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            tgtgcacatcactgactggtaccccactctcatttcactggctgaaggacagattgatga 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                          Sudduth-Klinger J;
                                                                                                                                                                                                                                                                                                                                                                                           C, Pot D, Kassam A;
n M, Drmanac S, Laba
Stache-Crain B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21.1%; Score 479; DB 21; Length 729; 99.8%; Pred. No. 3.9e-219;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 729 BP; 232 A; 178 C; 163 G; 143 T; 13 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 negative breast cancer, lung cancer, and colon cancer
                                                                                                                                                                                                                                                                                                                                                                        , Garcia PD, S
Kennedy GC, Pc
R, Dickson M,
                                                                                                                                                                                                                                                                                                                                                                                                                                            Jones LW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                     Escobedo J, Innis MA, G
Jese K, Randazzo F, Ken
manac R, Crkvenjakov R,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Claim 1; Page 924-925; 1097pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                            Garcia V,
                                                                                                                                                      98US-0085426.
98US-0085537.
98US-0085696.
98US-0105234.
                                                                                                               99WO-US10602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               529; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                            Leshkowitz D, Kita D,
                                                                                                                                                                                                                                                                                                                                                                                              Giese K,
Drmanac R,
                                                                                                                                                                                                                                                                                                      CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2000-126369/11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                    (CHIR ) CHIRON COR
(HYSE-) HYSEQ INC.
                                                                                                                                                                                                                                                                                                                                                                   Williams LT,
Reinhard C, G
Lamson G, Drm
                                                                                                                                                                                 15-MAY-1998;
15-MAY-1998;
21-OCT-1998;
27-OCT-1998;
                     WO9958675-A2
                                                                                                             13-MAY-1999;
                                                                                                                                                           14-MAY-1998;
                                                                 18-NOV-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1625
```

gaaattgcttacaggaaatcctggctacagcgactgggtccccctcagtcttcagcaa 1744

1685 227

ò a ð g ò g à 8 ð a

287

gaaattgettacaggaaateetggetacagegaetgggteeeeeeteagtettteageaa 346

```
Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase; ligase; hyperproliferative disorder; immunodeficiency disorder; autolimmune disorder; neurological disorder; metabolic disorder; inflammatory disorder; cardiovascular disorder; reproductive disorder; blood·related disorder; infectious disorder; gene therapy; cytostatic; anti arthritic; nephrotropic; anticoagulant; ss.
                                                                                                                                                                                               1924
                                                                                                                                                                                            cctgggaccgaaccggtggcacaatgaacggatcacctcgtcaactggcaaagtgtatg
                                       347 cctgggaccgaaccggtggcacaatgaacggatcaccttgtcaactggcaaaagtgtatg
                                                                                              gcttttcaacatcacagccgacccatatgagagggtggacctatctaacaggtatccagg
                                                                                                                                                gcttttcaacatcacagccgacccatatgagagggtggacctatctaacaggtatccagg
                                                                                                                                                                                                                                                                                                1925 gtatcccccaaagaccccagaagtaaccctaggctcaatggaggggtct 1974
                                                                                                                                                                                                                                                                                                                          cDNA encoding novel human enzyme polypeptide #421.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2000US-0209467.
2000US-0214886.
2000US-0215135.
2000US-0216647.
2000US-0216880.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAS41205 standard; cDNA; 629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2000US-0220964.
2000US-0224518.
2000US-0224519.
2000US-0225213.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2000US-0217487.
2000US-0217496.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             17-JAN-2001; 2001WO-US01239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2000US-0186350
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2000US-0189874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2000US-0190076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2000US-0198123.
2000US-0205515.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000US-0218290.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000US-0225214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20000S-0225757.
20000S-0225758.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2000US-0225759
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000US-0226279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 04-FEB-2000;
24-FEB-2000;
24-FEB-2000;
16-MAR-2000;
11-AMR-2000;
19-MAY-2000;
20-JUN-2000;
30-JUN-2000;
30-JUN-2000;
30-JUN-2000;
11-JUL-2000;
11-JUL-2000;
11-JUL-2000;
26-JUL-2000;
26-JU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       14-AUG-2000;
14-AUG-2000;
14-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                14-AUG-2000;
```

```
2000US-0231413
2000US-0231414
                                                                                                                                                                                                                     2000US-0232398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2000US-0241826
                                                                                                                      2000US-0230438
                                                                                                                                                                                                                                 2000US-0232399
                                                                                                                                                                                                                                                   2000US-0232401
                                                                                                                                                                                                                                                               2000US-0233063
                                                                                                                                                                                                                                                                       2000US-0233064
                                                                                                                                                                                                                                                                                           2000US-0234223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2000US-0241808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2000US-0246523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2000US-0246613
                                                                                                                                                                                                                                                                                                                       2000US-0234998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000US-0246475
                                                                                                                                                                                                                                                                                                                                           2000US-0235834
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0246477
                                                                                                                                                                                                                                                                                                                                                                                                               2000US-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2000US-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2000US-
22-AUG-2000;
22-AUG-2000;
23-AUG-2000;
23-AUG-2000;
30-AUG-2000;
01-SEP-2000;
01-SEP-2000;
01-SEP-2000;
                                                                            01-SEP-2000;
05-SEP-2000;
06-SEP-2000;
06-SEP-2000;
08-SEP-2000;
08-SEP-2000;
08-SEP-2000;
08-SEP-2000;
08-SEP-2000;
08-SEP-2000;
08-SEP-2000;
08-SEP-2000;
                                                                                                                                                                                                                                                                                                                     25-SEP-2000;
26-SEP-2000;
27-SEP-2000;
                                                                                                                                                                                                                                                                                                                                                 27-SEP-2000;
29-SEP-2000;
29-SEP-2000;
29-SEP-2000;
29-SEP-2000;
29-SEP-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                        02-0CT-2000;
02-0CT-2000;
13-0CT-2000;
20-0CT-2000;
20-0CT-2000;
20-0CT-2000;
20-0CT-2000;
20-0CT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20 - OCT - 2000
20 - OCT - 2000
11 - NOV - 2000
08 - NOV - 2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -NOV-2000;
-NOV-2000;
                                                                                                                                                                                                12-SEP-2000;
14-SEP-2000;
14-SEP-2000;
                                                                                                                                                                                                                               14-SEP-2000;
                                                                                                                                                                                                                                                                       14-SEP-2000;
                                                                                                                                                                                                                                                                                                   21-SEP-2000;
25-SEP-2000;
                                                                                                                                                                                                                                                                                                                                                                                                             02-OCT-2000;
                                                                                                                                                                                                                                        14-SEP-2000;
                                                                                                                                                                                                                                                             14-SEP-2000;
                                                                                                                                                                                                                                                                               14-SEP-2000;
                                                                                                                                                                                                                                                                                           -SEP-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                     02-OCT-2000;
02-OCT-2000;
```

```
The present invention relates to the isolation of novel human enzyme polypeptides (AAU22915-AAU3814), and the CDNA and genomic sequences encoding them. The enzyme polypeptides of the invention may comprise the functional classes of oxidoreductases, transferases, hydrolases, lyases, isomerases or ligases. The sequences of the invention are useful in the disorders including hyperproliferative disorders (e.g. cancer), immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g. athritis), enurological disorders (e.g. AILAHEMER'S disease), metabolic disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis), blood-related disorders (e.g. haemophilia), reproductive disorders (e.g. infertility) and infectious disorders (e.g. Influenza). The polymoclocides of the invention can also be used in gene therapy. AAS40785-AAS41684 represent CDNA sequences encoding for the novel human enzyme polypeptides of the invention can also be used in gene therapy. AAS41684 represent cDNA sequences encoding for the novel human enzyme polypeptides of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Novel polypeptides and polynucleotides useful for diagnosing, preventing, treating neural, immune system, muscular, reproductive, pulmonary, cardiovascular, renal, proliferative disorders and cancerous diseases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 4; SEQ ID No 431; 1180pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ruben SM
                                                                                                                                                                                                                                                                                                                                                                     2000US-0251856.
2000US-0251868.
2000US-0251869.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (HUMA-) HUMAN GENOME SCI INC
                                                                                    2000US-0249213
                                                                                                                   2000US-0249215
                                                                                                                                                            2000US-0249218
                                                                                                                                                                                          2000US-0249245
                                                                                                                                                                                                                       2000US-0249265
                                                                                                                                                                                                                                                               2000US-0249300
2000US-0250160
                                                                                                                                                                                                                                                                                                                           2000US-0251988.
2000US-0256719.
                                                                                                                                                                                                                                                                                                                                                        2000US-0251479
                                                                                                                                                                                                                                     2000US-0249297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Rosen CA, Barash SC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2001-465566/50.
P-PSDB; AAU23335.
                       17 - NOV - 2000;
                                                                                                                                                                                       17-NOV-2000;
17-NOV-2000;
17-NOV-2000;
17-NOV-2000;
17-NOV-2000;
01-DEC-2000;
                                                                                                                                                                                                                                                                                           01-DEC-2000;
05-DEC-2000;
05-DEC-2000;
06-DEC-2000;
08-DEC-2000;
08-DEC-2000;
08-DEC-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                               08-DEC-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          diseases
```

1;

Gaps

1;

Indels

ö

0; Mismatches

Query Match 19.2' Best Local Similarity 99.8 Matches 554; Conservative

19.2%; Score 434; DB 22; Length 629; 99.8%; Pred. No. 1.4e-197;

BP; 189 A; 151 C; 124 G; 162 T; 3 other;

Sequence 629

ftp.wipo.int/pub/published\_pct\_sequences.

```
(CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                  RESULT
ò
                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                   gcctgggactatgacaatggcatatactccacacagatgtacactcagagagtacagcaa 1067
                                                                                                                                                                                         atcttagcttcccataaccccacaaagcctatatttttatatattgcctatcaagctgtt 1127
                                                                                                                   888
                                 ttattactggaaagtatcagatacaccggacttcaacattctatcataagacctaccc 768
               73
                                                                                                                                                                                                                                                                                       Human colon cancer cell line polynucleotide sequence SEQ ID NO:269
          Human; colon cancer; tumour; diagnosis; gene expression product; probe; detection; cancerous state; metastasis; identification; breast cancer; oestrogen receptor-positive breast cancer; therapy; oestrogen receptor-negative breast cancer; lung cancer; ss.
    gagatgtgggttaccacggatctgagattaaaacacctactcttgacaagctcgctgccg
                          aaggagttaaactggagaactactatgtccagcctatttgcacaccatccaggagtcagt
                                                      BP
                                                                                                                                                                                                                                                                                                                                                                            98US-0085426.
98US-0085537.
98US-0085696.
98US-0105234.
                                                                                                                                                                                                                                                      AAA00278 standard; cDNA; 294
                                                                                                                                                                                                                                                                                                                                                                  99WO-US10602
                                                                                                                                                                                                                                                                            19-MAY-2000 (first entry)
                                                                                                                                                                                                               cattcaccactgcaa 1142
                                                                                                                                                                                                                    (CHIR ) CHIRON CORP. (HYSE-) HYSEQ INC.
                                                                                                                                                                                                                                                                                                                                Homo sapiens
                                                                                                                                                                                                                                                                                                                                           WO9958675-A2
                                                                                                                                                                                                                                                                                                                                                                            14-MAY-1998;
15-MAY-1998;
15-MAY-1998;
21-OCT-1998;
27-OCT-1998;
                                                                                                                                                                                                                                                                                                                                                                  13-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                       18-NOV-1999.
                                                                                                                                                                                                                                                                 AAA00278;
                                                                                                                                                                                                                                           1
                                                                                                                                                                              434
                                                                                  194
                                                                                                                                                                  1008
                                                                                                                                                                                                               1128
                                                                                                                                                                                                                         554
   589
                          649
                                                 709
                                                                        694
                                                                                              829
                                                                                                                                                                                         1068
                                                                                                                                                                                                                                                 AAA00278
                                                ò
                                                           a
                                                                                 g
                                                                                              ò
                                                                                                        qq
                                                                                                                               g
                                                                                                                                           ô
                                                                                                                                                                            g
                                                                                                                                                                                                   g
                                                                                                                                                                                                               ò
                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                           ò
              g
                          õ
                                     g
                                                                        ö
                                                                                                                    ö
                                                                                                                                                      a
                                                                                                                                                                 ò
                                                                                                                                                                                        õ
```

```
The AAA0010 to AAA02716 represent polynucleotides isolated from cDNA libraries constructed from human colon cancer cell lines. The present invention also describes a method of detecting differentially expressed genes correlated with a cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample derived from a cell suspected of being cancerous, where detection of the differentially expressed gene product is correlated with a cancerous state of the cell from which the test sample was derived. The polynucleotides sequences can be used in a method for detecting differentially expressed genes correlated with a cancerous state of a mammalian cell. The polynucleotides can also be used as probes for detecting and mapping related genes. They can be used in diagnosis and cancering and disorders (e.g. identification of prognosis of diseases and disorders (e.g. identification of cancer, or responsiveness of cancer to therapy). This is particularly for breast cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-very
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                    mammalian
                     Randazzo F, Kennedy GC, Pot D, Kassam A;
Crkvenjakov R, Dickson M, Drmanac S, Labat
Garcia V, Jones LW, Stache-Crain B;
     ņ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Human colon cancer cell line polynucleotide sequence SEQ ID NO:2361.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Human; colon cancer; tumour; diagnosis; gene expression product; probe; detection; cancerous state; metastasis; identification; breast cancer; oestrogen receptor-positive breast cancer; therapy; oestrogen receptor-negative breast cancer; lung cancer; ss.
Sudduth-Klinger
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
                                                                                                                                                                                    determine cancerous states of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 294;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 294 BP; 79 A; 46 C; 51 G; 102 T; 16 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 21; L
5.4e-16;
hes 0;
Garcia PD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2.4%; Score 55; DB
100.0%; Pred. No. 5.4
ive 0; Mismatches
                                                                                                                                                                                                                                                              Claim 1; Page 244-245; 1097pp; English.
Innis MA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BP.
                                                                                                                                                                                 Polynucleotide library used to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAA02370 standard; cDNA; 1082
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98US-0085426.
98US-0085537.
98US-0085696.
98US-0105234.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99WO-US10602
Escobedo J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 100.C
Matches 55; Conservative
                                                                               Ď
                           Giese K,
                                                   Drmanac R,
                                                                                                                             WPI; 2000-126369/11.
                                                                               Kita
                                                                               Leshkowitz D,
Williams LT,
Reinhard C, (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              W09958675-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     15-MAY-1998;
21-OCT-1998;
27-OCT-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19-MAY-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18-NOV-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    14-MAY-1998
15-MAY-1998
                                                         Lamson G,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAA02370;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           12
                                                                                                                                                                                                            cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAA02370
```

```
98US-0079663.
98US-0079664.
98US-0079689.
98US-0079728.
                                                                                                                                                                                              98US-0080194.
98US-0080327.
98US-0080338.
98US-0080334.
98US-0081049.
98US-00811071.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98US-0083499
98US-0083500
98US-0083545
98US-0083554
98US-0083558
98US-0083559
                                                                                                                                                 98US-0079923.
98US-0080105.
98US-0080107.
98US-0080165.
                                                                                                                                                                                                                                                                                                             98US-0081229.
98US-0081817.
98US-0081838.
98US-0081952.
                                                                                                                                                                                                                                                                                                                                                                                 98US-0082569.
98US-0082700.
98US-0082704.
                                                                                                                                                                                                                                                                                                                                                                                                                                        98US-0083336.
98US-0083336.
98US-0083322.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98US-0083495.
98US-0083495.
98US-0083496.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98US-0084414.
98US-0084411.
98US-0084598.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0084600.
98US-0084627.
98US-0084637.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98US-0085704.
98US-0085704.
98US-0086023.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98US-0084640.
98US-0084643.
                                                         98US-0079294
98US-0079656
                                                                                                                                        98US-0079920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98US-0085580
98US-0085582
                                                                                                                                                                                                                                                                                                                                                             98US-0081955
                                                                                                                                                                                                                                                                                                                                                                                                                             98US-0082767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98US-0085338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0085573
                                                                                                                                                                                                                                                                                                                                                                                                                    98US-0082804
                                 20-MAR-1998;
20-MAR-1998;
25-MAR-1998;
26-MAR-1998;
27-MAR-1998;
                                                                                                               27 - MAR - 1998;
27 - MAR - 1998;
30 - MAR - 1998;
30 - MAR - 1998;
31 - MAR - 1998;
                                                                                                                                                                       31-MAR-1998;
31-MAR-1998;
31-MAR-1998;
01-APR-1998;
01-APR-1998;
01-APR-1998;
01-APR-1998;
                                                                                          27-MAR-1998;
27-MAR-1998;
                                                                                                                                                                                                                                                                                       09-APR-1998;
09-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 29-APR-1998;
29-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        05-MAY-1998;
06-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             28-APR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29-APR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-MAY-1998
  ö
                                                                                                                                                                                         penes correlated with a cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test asmaple derived from a cell suspected of being cancerous, where detection of the differentially expressed gene product is correlated with a cancerous state of the cell from which the test sample was derived. The polynucleotides sequences can be used in a method for detecting differentially expressed genes correlated with a cancerous state of a mammalian cell. The polynucleotides can also be used as probes for detecting and mapping related genes. They can be used in diagnosis and prognosis of diseases and disorders (e.g. identification of pre-metastatic or metastatic cancerous states, stages of cancer, or responsiveness of cancer to therapy). This is particularly for breast cancer, cestrogen receptor-positive breast cancer, cestrogen receptor-
                                                                                                                                                          AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
libraries constructed from human colon cancer cell lines. The present
invention also describes a method of detecting differentially expressed
                     Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J; Reinhard C, Glase K, Randazzo F, Kennedy GC, Pot D, Kassam A; Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I; Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
probe; blood coagulation disorder; cancer; cellular adhesion disorder;
secreted protein; transmembrane protein; ss.
                                                                                                     Polynucleotide library used to determine cancerous states of mammalian
                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                              .
0
                                                                                                                                                                                                                                                                                                                                                                                                    1.9%; Score 44; DB 21; Length 1082;
100.0%; Pred. No. 1e-10;
/ative 0; Mismatches 0; Indels (
                                                                                                                                                                                                                                                                                                                                                                     Sequence 1082 BP; 230 A; 182 C; 155 G; 190 T; 325 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                negative breast cancer, lung cancer, and colon cancer.
                                                                                                                                     Claim 1; Page 938; 1097pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Human PRO708 hybridisation probe.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98US-0077450.
98US-0077632.
98US-0077641.
98US-0077649.
98US-0077791.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99WO-US05028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAZ33990 standard; DNA; 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             07-DEC-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                              WPI; 2000-126369/11.
                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
(HYSE-) HYSEQ INC.
                                                                                                                                                                                                                                                                                                                                                                                                        Ma.
Local S...
44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11-MAR-1998;
11-MAR-1998;
11-MAR-1998;
12-MAR-1998;
13-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO9946281-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              16-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAZ33990;
                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                 cells
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                  Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAZ33990
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
```

W, Gerber H, Gerritsen ME; Gurney AL, Hillan KJ; Paoni NF, Roy MA;

Wood WI

Eaton DL;

Desnoyers L,

New

```
sequence tag) sequences which produces corrected or transmembrane PRO polypeptides and polypeptides have cytostatic activity. The PRO polynucleotides and polypeptides have cytostatic the presence of PRO polypeptides in samples, for linking bloactive molecules to cells and for modulating biological activities of cells, using the polypeptides for specific targeting. The polypeptide for specific targeting or the treatment of cancers to perpeptide pairs provide specific targeting of bloactive molecules to cells. AAC78600 to AAC7867 represent PCR primers and probes used in the isolation of the PRO polynucleotide sequences.
                                                                                                                                                                                                                                                                                                                                                                                                              Novel PRO polypeptides and polynucleotides used in detection methods, to target bioactive molecules to specific cells, and to modulate cellular activities -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Human; immune; haematopoletic; immune/haematopoletic antigen; cancer; cytostatic; gene therapy; vaccine; metastasis; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:8808
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAC78458 to AAC78599 represent polynucleotide and EST (expressed
                                                                                                                                                                                                                                                                 KP, Botstein D, Desnoyers I
E, Fong S, Gao W, Gerber
PJ, Grimaldi CJ, Gurney AL,
Napier MA, Pan J, Paoni NE
TA, Tumas D, Williams PM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1.4%; Score 32; DB 21; L
ilarity 100.0%; Pred. No. 5.6e-05;
Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 53 BP; 14 A; 13 C; 14 G; 11 T; 1 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            795 gccaccctacctcagaaactgaaggaggttgg 826
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 gccaccctacctcagaaactgaaggaggttgg 32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Example 19; Page 250; 636pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ВР
                                                                99WO-US28551.
99WO-US28555.
99WO-US30095.
99WO-US31243.
2000WO-US0219.
2000WO-US0277.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAK63748 standard; cDNA; 774
                99US-0145698.
99US-0162506.
99WO-US28313.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          17-JAN-2001; 2001WO-US01354.
                                                                                                                                                                                                                                                                 Ashkenazi AJ, Baker KP, B
Ferrara N, Filvaroff E, F
Goddard A, Godowski PJ, G
Kljavin IJ, Kuo SS, No Je
Shelton DL, Stewart TA, T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                  (GETH ) GENENTECH INC.
                                                                                                                                                                                                                                                                                                                                                                           WPI; 2000-611443/58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity
hes 32; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200157182-A2
              26-JUL-1999;
29-OCT-1999;
30-OCT-1999;
02-DEC-1999;
16-DEC-1999;
30-DEC-1999;
30-DEC-1999;
30-DEC-1999;
66-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens.
                                                                                                                                                                                               06-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             06-NOV-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAK63748;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAK63748
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                              present invention describes secreted and transmembrane polypeptides
                                                                                                                                                                                                                                                                                                                                                                                                          and their polynucleotides. The nucleotide sequences are useful as sources of probes, primers, for chromosome mapping, and for generation of antisense sequences. They can also be used to create transgenic animals. The proteins can be used to treat a variety of diseases and disorders, depending on their function. Diseases that may be treated include blood coagulation disorders, cancers and cellular adhesion disorders. They may also be used to raise antibodies. AAZ33891 to AAZ34388, and AAX41685 to AAX41774 represent polynucleotide and polypeptide sequence given in the exemplification of the present
                                                                                                                                                                                                                                                                                    secreted and transmembrane polypeptides and their polynucleotides, ful for treating blood coagulation disorders, cancers and cellular
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Human; secreted protein; transmembrane protein; PRO; EST; cytostatic; expressed sequence tag; detection; cancer; PCR primer; probe; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                 Chen J;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1.4%; Score 32; DB 20; Length 53
100.0%; Pred. No. 5.6e-05;
ative 0; Mismatches 0; Indels
                                                                                                                                                                                                                 Baker KP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 53 BP; 14 A; 13 C; 14 G; 11 T; 1 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Human PRO708 hybridisation probe SEQ ID NO:117.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Goddard A, Gurney A, Yuan J,
                                                                                                                                                                                                                                                                                                                                                          Example 19; Page 197; 530pp; English.
              98US-0086414.
98US-0086430.
98US-0086486.
98US-0087098.
98US-0087106.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99WO-US05028.
99US-0123957.
99US-0126773.
99US-0130232.
99US-0131445.
98US-0086392
                                                                                                                      98US-0094651
                                                                                                                                           98US-0100038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000WO-US04341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-FEB-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity 100.
les 32; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAC78675 standard; DNA;
                                                                                                                                                                            (GETH ) GENENTECH INC
                                                                                                                                                                                                                                                 WPI; 1999-551358/46.
                                                                                                                                                                                                                                                                                                                       adhesion disorders
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200053756-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18-FEB-2000;
            22-MAY-1998;
22-MAY-1998;
22-MAY-1998;
28-MAY-1998;
28-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  29-MAR-1999;
21-APR-1999;
28-APR-1999;
14-MAY-1999;
                                                                                                                      30-JUL-1998;
11-SEP-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   08-MAR-1999;
                                                                                                       28-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              14-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12-MAR-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              nvention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAC78675;
                                                                                                                                                                                                             Wood WI,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    195
                                                                                                                                                                                                                                                                                                                                                                                            The
```

ö

Gaps

;

Indels

, 0

AAC7867

à a Length 53;

```
2000US-0190076.
2000US-0198123.
2000US-0205515.
2000US-020467.
2000US-0214886.
2000US-0215135.
                                                                                                                                        2000US-0216880.
2000US-0217487.
2000US-0217496.
2000US-0218290.
2000US-0220963.
2000US-0220963.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2000US-0231413.
2000US-0231414.
2000US-0232080.
2000US-0232081.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2000US-0231968.
2000US-0232397.
2000US-0232398.
                               2000US-0184664.
2000US-0186350.
2000US-0189874.
                                                                                                                                                                                                                                                                                              20000S-0225447.
20000S-0225757.
20000S-0225758.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000US-0231242.
2000US-0231243.
2000US-0231244.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2000US-0232400.
2000US-0232401.
2000US-0233063.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2000US-0233064.
2000US-0233065.
2000US-0234223.
                                                                                                                                                                                                                    2000US-0224519.
2000US-0225213.
2000US-0225214.
                                                                                                                                                                                                                                                               2000US-0225267.
2000US-0225268.
2000US-0225268.
                                                                                                                                                                                                                                                                                                                                                                                                                     2000US-0229343.
2000US-0229344.
2000US-0229345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    2000US-0229509.
2000US-0229513.
2000US-0230437.
          2000US-0179065
2000US-0180628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2000US-0230438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2000US-0232399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2000US-0234997
2000US-0234998
                                                                                                                                                                                                                                                     2000US-0225266
                                                                                                                                                                                                                                                                                                                                                     2000US-0226681
2000US-0226868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2000US-0234274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2000US-0236370
                                                                                                                                                                                                                                                                                                                                                                            2000US-0227182
                                                                                                                                                                                                                                                                                                                                                                                                2000US-0228924
                                                                                                                                                                                                                                                                                                                                                                                                            2000US-0229287
                                                                                                                                                                                                                                                                                                                                             2000US-022627
        131-JAN-2000)
44-FEB-2000)
42-FEB-2000)
16-MAR-2000)
16-MAR-2000)
18-MAR-2000)
19-MAY-2000)
19-MAY-2000)
19-MAY-2000)
10-JUL-2000)
11-JUL-2000)
11-J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29-SEP-2000;
29-SEP-2000;
```

PR 02-0CT-2000; 2000US-0236802.
PR 02-0CT-2000; 2000US-0237031.
PR 02-0CT-2000; 2000US-0237031.
PR 13-0CT-2000; 2000US-0237031.
PR 13-0CT-2000; 2000US-0237031.
PR 20-0CT-2000; 2000US-023937.
PR 20-0CT-2000; 2000US-023937.
PR 20-0CT-2000; 2000US-0241785.
PR 20-0CT-2000; 2000US-0241785.
PR 20-0CT-2000; 2000US-024186.
PR 20-0CT-2000; 2000US-024186.
PR 20-0CT-2000; 2000US-0241876.
PR 20-0CT-2000; 2000US-0241876.
PR 20-0CT-2000; 2000US-0241876.
PR 20-0CT-2000; 2000US-024617.
PR 20-0CT-2000; 2000US-024622.
PR 20-0CT-2000; 2000US-024920.
PR 20-0CT-2000; 2000U

.X VT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,

```
useful for preventing, diagnosing and/or treating cancers and
metastasis -
```

Claim 1; SEQ ID NO 8808; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoletic antigen (I)
amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
activity, and can be used in gene therapy and vaccine production. (I)
proteins and polynucleotides may be used in the prevention, diagnosis and
treatment of diseases associated with inappropriate (I) expression. For
example, they may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patient's genome
that affect the activity of (I) by expressing inactive proteins or to
supplement the patients own production of (I). Additionally, (I)
complement the patients own production of (I). Additionally, (I)
coplynucleotides may be used to produce the secreted (I), by inserting
the nucleic acids into a host cell and culturing the cell to express the
protein. (I) proteins and polynucleotides may be used to prevent.

Concers and cancer metastases of haematopoietic-derived cells. AAK64703
concers and cancer metastases of haematopoietic antigen genomic
sequences from the present invention. AAK54942 to AAK89690 and AAM82169
crepresent sequences used in the exemplification of the present invention.

Sequence 774 BP; 215 A; 194 C; 189 G; 173 T; 3 other;

; 0 Gaps .; 0 Query Match 1.1%; Score 26; DB 22; Length 774; Best Local Similarity 100.0%; Pred. No. 0.043; Matches 26; Conservative 0; Mismatches 0; Indels

õ

q

Search completed: July 23, 2002, 19:01:42 Job time: 11368 sec

```
ACC1342 Home Sapi
ACC1342 Home Sapi
ACC01342 Home Sapi
ACL05414 Home Sapi
ACL05343 Home Sapi
ACC05383 Home Sapi
ACC091322 Mus muscu
ACC09604 Rattus no
BCC009528 Home Sapi
BCC012659 Mus muscu
BCC009528 Home Sapi
ACC44865 Home Sapi
ACC44865 Home Sapi
ACC44865 Home Sapi
ACC04865 Home Sapi
ACC01833 Home Sapi
ACC01426 Mus muscu
BCC021426 Mus muscu
BCC021426 Home Sapi
ACC068939 Home Sapi
                                                                                                                                                                                                         ACO68939 Homo Sapir
ACO91369 Rattus no
AC106436 Rattus no
AC101256 Homo Sapi
AC011256 Homo Sapi
AC0713548 Homo Sapi
AC073548 Homo Sapi
AC073548 Homo Sapi
AC072815 Homo Sapi
AC0702815 Homo Sapi
AC016770 Homo Sapi
AC177154 Sequence
AX277154 Sequence
AX277154 Sequence
AX277159 Kenopus 1
BC015055 Homo Sapi
AC015055 Homo Sapi
AC015058 Homo Sapi
AC015058 Homo Sapi
AC015058 Homo Sapi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PAT 30-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                       BC021490 Mus muscu
AC012851 Drosophil
AL671136 Human DNA
                  Homo sapi
Homo sapi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 (bases 1 to 2253)
Glucksmann,M.A., Williamson,M., Rudolph-Owen,L.A. and Tsai,F.Y.
Human sulfatases
Patent: WO 015541-A 8 02-AUG-2001;
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
          AX206973
AC104779
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1. .2253
/organism="Homo sapiens"
/organism="taxon:9606"
324. .2123
/note="unnamed protein product"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          · DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 8 from Patent W00155411. AX206967.1 GI:15394724
                                      AC013692
AC105414
AK075414
AK075411
AC105383
AC091322
AC091322
BC009559
BC012659
AC108339
                                                                                                                                      CBRG36E21
AC044865
AP001838
                                                                                                                                                                                                                               AC106436
AC008567
AC011256
                                                                                                                                                                                                                                                                               AP002386
AP002815
AC097993
AC011451
                                                                                                                                                                                                                                                            AP001142
AC073548
                                                                                                                                                                                        BC021427
                                                                                                                                                                    RATARSB
                                                                                                                                                                                                                                                                                                                                                                                                               AX349119
                                                                                                                                                                                                                      AC091369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /codon_start=1
                                                                                                                   20
                                                                                                                                                            5100
                                                                                                                                                                                                                                                                                                                                                               200
                                                                                                                                                                                                                               153400
153704
153882
161591
170537
174484
174484
176852
180672
191919
2253
152880
1800
165147
165147
175143
175143
175141
175141
122992
2003
28735
138866
131654
                                                                                                                                                                   1603
2491
3695
118722
136503
                                                                                                                                                                                                                                                                                                                                                                                                              4858
4918
17514
20280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           human.
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
AX206967
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                                                            LOCUS
                    000
                                                                           (without alignments)
11580.936 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                             1 cacgcgtccgcccacgcgtc......tgccacctggtgccgaattc 2266
                                                                 July 23, 2002, 15:12:04; Search time 4094.62 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Description
                                                                                                                                                                                                                      3595312
        GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                             1797656 segs, 10463268293 residues
                                                                                                                                                                                                                       parameters:
                                                                                                                                                                                                                                                                      summaries
                                             nucleic search, using sw model
                                                                                                                                                OLIGO_NUC
Gapop 60.0 , Gapext 60.0
                                                                                                                                                                                                                    Total number of hits satisfying chosen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                em_vi:*
em_htg_hum:*
em_htg_inv:*
em_htg_other:*
em_htg_other:*
                                                                                                                                                                                                                                                                      Post-processing: Listing first 45
                                                                                                                                                                                                                                        length: 0
length: 2000000000
                                                                                                       US-09-495-823-8
2266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB
                                                                                                                                                                                                                                                                                                 9b_ba:*
9b_htg:*
9b_on:*
9b_on:*
9b_ov:*
9b_pt:*
9b_pt:*
9b_pt:*
9b_pt:*
9b_pt:*
9b_re:*
9b_re:*
9b_re:*
9b_re:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                    em_fun:*
em_hum:*
em_in:*
em_au:*
em_au:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       em_pat:*
em_ph:*
em_pl:*
em_ro:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                em_sts:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     em_or:*
em_ov:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         em_un:*
                                                                                                                                                                                                                                                                                         GenEmbl:*
                                                                                                        Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                       Minimum DB seq
Maximum DB seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score
                                                                                                                                                Scoring table:
                                             OM nucleic
                                                                                                                                                                                                 Word size
                                                                                                                                                                              Searched:
                                                                                                                                                                                                                                                                                         Database
                                                                  eo
::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Result
```

Wed

```
/protein_id="CAC60191.1"

/db_xref="GI-15394725"

/translation="WaPRGCAGHPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYL
SWGQALEEBEEGALLAQAGKLEPSTYSTSQPHLIFILADDQGFRDVG*HGSEIKTPT
SWGQALEEBEEGALLAQAGKLEPSRSQFTTGKYQIHTCLLTYGYL
SWGQALEEBEEGALLAQAGKKLEPSTYSTSQPHLIFILADDQGFRDVG*HGSEIKTPT
LDKLAAEGYKLENYYQPICTPSRSQFTTGKYQIHTCHQASITRPTQPACLPLDNATL
PQKLKEVGYSTHWYGWHLGFYRKEMPTRRGPDTFFGSLLGSGDYTHYKCDSPGMC
GYDLYENDNAAWDYDNGITSTYGYTQYTQCHGASHNPTKPIFLYTATYGFYNHYKCDSPGMC
GYDLYGYTYWGGTYWGGTRAVGYGYHSPLLKNGTYCKELVHITDWYPTLISLAGGOT
GGSNWPRGSKGTYWGGTRAVGYHSPLLKNGTYCKELVHITDWYPTLISLAGGOT
DEDLQLGYTYGTYGTRASGNYGTAVPRYPRDATTLISLAGGOT
DEDLQLGYTWLLTGNPGTSNUCHNNHHERITLSTGKSVMLFNITADPYERVDL
SNRYPGTYKKLLTGNPGTSNUTHYPPRYPRDPRSNPRLNGGVWGPWYKEETKKKRPSK
NQHWKLLTGNPGTSNUTHYPPRYPRYPRDPRSNPRLNGGVWGPWYKEETKKKRPSK
NQHWKLGYGKKKKKKOOKAYGSTCHSGCYTGC"
                                                                                                                                           ·-
                                                                                                                                                                                     180
                                                                                                                                                                                                                                 180
                                                                                                                                                                                                                                                240
                                                                                                                                                                                                                                                              240
                                                                                                                                                                                                                                                                             300
                                                                                                                                                                                                                                                                                           300
                                                                                                                                                                                                                                                                                                          360
                                                                                                                                                                                                                                                                                                                       360
                                                                                                                                                                                                                                                                                                                                       420
                                                                                                                                                                                                                                                                                                                                                     420
                                                                                                                                                                                                                                                                                                                                                                    480
                                                                                                                                                                                                                                                                                                                                                                                  480
                                                                                                                                                                                                                                                                                                                                                                                                 540
                                                                                                                                                                                                                                                                                                                                                                                                               540
                                                                                                                                                                                                                                                                                                                                                                                                                              009
                                                                                                                                                                                                                                                                                                                                                                                                                                            900
                                                                                                                                                                                                                                                                                                                                                                                                                                                          099
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       720
                                                                                                                                            Gaps
                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    tetggateetetgeeteeteaettatggttaeetgteetggggeeaggeettagaagagg
                                                                                                                                           1;
                                                                                                                                            Indels
                                                                                                                              Length
                                                                                                                                           5;
                                                                                                                              . 9
                                                                                                                              DB
                                                                                                                             Score 1877; DI
Pred. No. 0;
0; Mismatches
                                                                                                                                           ·;
                                                                                                                             82.8%;
ilarity 99.7%;
Conservative
                                                                                                                                     lmilarity
                                                                                                  ď
                                                                                                 624
                                                                                                                                           2247;
                                                                                                                                     S
                                                                                                                              / Match
Local S
                                                                                                BASE COUNT
ORIGIN
                                                                                                                              Query Ma
Best Loca
Matches
                                                                                                                                                          ч
                                                                                                                                                                       ч
                                                                                                                                                                                       61
                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                             541
                                                                                                                                                                                                                                                                                                                                                                                                                                           541
                                                                                                                                                                                                                                 121
                                                                                                                                                                                                                                                                             241
                                                                                                                                                                                                                                                                                           241
                                                                                                                                                                                                                                                                                                         301
                                                                                                                                                                                                                                                                                                                       301
                                                                                                                                                                                                                                                                                                                                       361
                                                                                                                                                                                                                                                                                                                                                     361
                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                  421
                                                                                                                                                                                                                                                                                                                                                                                                 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                          601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         601
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     561
                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    721
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 721
                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                      Ω
                                                                                                                                                                                                   g
                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                        ő
                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                          ö
                                                                                                                                                                                      à
                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                          QQ
                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

1800 1140 1140 1260 1560 1680 1920 1020 1020 1080 1080 1560 960 840 900 960 tatggcttttcaacatcacagccgacccatatgagagggtggacctatctaacaggtatc caggaategtgaagaageteetaeggaageteteaeagtteaaaaetgeagtgeegg Lacctcagaaactgaaggaggttggatattcaacgcata ttgatacctttttttggttcccttttgggaagtggggattactatacacactacaatgtg gcaacctgggaccgaaccggtggcacaatgaacggatcacctcgtcaactggcaaaagtg tacctc 1141 1381 1441 1501 1741 1021 1081 1081 1141 1201 1261 1261 1321 1381 1441 1501 1561 1621 1681 1741 1801 1861 781 781 841 841 901 901 961 961 1021 1201 1321 1561 1621 1681 1801 g ´ qq g qq qq g Q qq g g Ω Op g qq ò g qq οy g οy ò δy ογ ŏ οy οy οy οy δy δ δ ò δ qq á δ

ATCCTAGCGGATGATCAGGGATTAGAGATGTGGGTTACCACGGATCTGAGATTAAAACA 300	cctactcttgacaagctcgctgccgaaggagttaaactggagaactactatgtccagcct 683 	atttgcacaccatccaggagtcagtttattactggaaagtatcagatacacaccggactt 743 	caacattctatcataagacctacccaacccattttacctctggacaatgccaccta 803	cctcagaaactgaaggaggttggatattcaacgcatatggtcggaaatggcacttgggt 863 	ttttacagaaaagaatgcatgcccaccagaagagtttgatacctttttggttccctt 923 	ttgggaagtgggattactatacacactacaaatgtgacagtcctgggatgtgtggctat 983 	gacttgtatgaaaacgacaatgctgcctgggactatgacaatggcatatactccacaca 1043 	atgtacactcagagagtacagcaaatcttagcttcccataaccccacaaagcctatattt 1103 	ttatatattgcctatcagctgttcattcaccactgcaagctcctggcaggtatttcgaa 1163 	cactaccgatccattatcaacataaacaggaggagatatgctgccatgcttcctgctta 1223 	gatgaagcaatcaacaacgtgacattggctctaaagacttatggtttctataacaacagc 1283 	attatcatttactcttcagataatggtggccagcctacggcaggaggagtaactggcct 1343 	ctcagaggtagcaaaggaacatattgggaaggagtccgggctgtaggctttgtgcat 1403 	agcccacttctgaaaacaagggaacagtgtgtaaggaacctgtgcacatcactgactg	taccccattctcatttcactggctgaaggacagattgatgaggacattcaactagatggc 1523 	tatgatatctgggagaccataagtgaggtcttcgctcaccccgagtagatattttgcat 1583 	aacattgaccccatatacaccaaggcaaaaatggctcctgggcagcaggctatgggatc 1643 	tggaacactgcaatccagtcagccatcagagtgcagcactggaaattgcttacaggaaat 1703
Db 241	Oy 624 Db 301	Qy 684 Db 361	Qy 744 Db 421	Oy 804 Db 481	Qy 864 Db 541	Oy 924 Db 601	Qy 984 Db 661	Oy 1044 Db 721	Qy 1104 Db 781	Qy 1164 Db 841	Qy 1224 Db 901	Oy 1284 Db 961	Qy 1344 Db 1021	Oy 1404 Db 1081	Oy 1464 Db 1141	Qy 1524 Db 1201	Qy 1584 Db 1261	Qy 1644
Db 1861 CAGGAATCGTGAAGAAGCTCCTACGGAGGCTCTCACAGTTCAACAAAACTGCAGTGCCGG 1920	QY 1921 tcaggtatcccccaaagaccccagaagtaaccctaggctcaatggagggtctagggac 1980 	QY 1981 catggtatagagagaaaccaagaaaaagagccaagcaaaatcaggctgagaaaagc 2040 	Oy 2041 aaaagaaaagc-aaaaaagaagaagaagaagcagcaggtccaggttcaacttgcc 2099	Oy 2100 attcaggtgttacttgtgggtaagcacaaatatttcctgtttggttaaactttaatcagt 2159	Oy 2160 tcttatctttcatctgtttcctaggtaaaccagcaaatttggctcgataatatcgctggc 2219	Oy 2220 ctaagcgtcaggcttgtttcatgctgtgccac 2252 	OC BAG VECASI AND	ITION Sequence 14 from Patent W00155411. SION AX206973 ON AX206973.1 GI:15394729	SOURCE human. ORGANISM Homo sapiens Craniata; Vertebrata; Euteleostomi; Nammalia: Futboria, neimatoc, Craniata; Vertebrata; Euteleostomi;	, Rudolph-Owen,L.A. a	Ç	Ado xref="taxon:9606"   Ado xref="taxon:9606"   Ado xref="taxon:9606"   Ado taxon:9608   Ado taxon:9608	Query Match Best Local Similarity 99.9%; Pred. No. 0; Matches 1649; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	atgottcccaggggctgtgcggggcatccgccttctccacaggcctgtctgt	384 cctggaaagatgctagcaatggggggctggcaggattctggatcctctgcctcctcact	444 tatggttacctgtcctgggccaggccttagaagaggagaagaaggggccttactagct 	504 caagetggagagaactagagcccagcacaacttccacctcccagccccattttc	564 atcctagcggatgatcagggatttagagatgtgggttaccacggatctgagattaaaaca

	AGAGGAACCCTAGGCTCAATGGAGGGTCT 1651  AGAAGTAACCCTAGGCTCAATGGAGGGGTCT 1651  AC104779  AC	2 (bases 1 to 152861) Waterston, R.H. Direct Submission Submitted (21-DEC-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA 3 (bases 1 to 152861) Waterston, R.H. Direct Submission Submitted (09-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA On Feb 9, 2002 this sequence version replaced gi:18482303.	ø	/clone="RP11.26P13" SE COUNT 48870 a 27681 c 28505 g 47805 t IGIN Query Match 53.3%; Score 1207; DB 9; Length 152861; Best Local Similarity 99.7%; Pred. No. 0; Matches 1527; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
0y 1704 0y 1704 0y 1764 0y 1764 0y 1824 0y 1884 0y 1884 0y 1984	D 1621 DD 1621 AC104779/C LOCUS LOCUS LOCUS ACCESSION VERSION VERSION VERSION VERSION VERVORDS SOURCE ORGANISM AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL AUTHORS TITLE JOURNAL	FEATURES	DRIGIN Query Match Best Local

```
10137
              10077
                        10017
                                                               9777
                                                                                                      9537
                                                                                                            1441
                                                                                                                      1501
                                                                                                                           9417
                                  9957
                                       1021
                                            9897
                                                 1081
                                                      9837
                                                           1141
                                                                    1201
                                                                          9717
                                                                               1261
                                                                                   9657
                                                                                        1321
                                                                                             9597
                                                                                                  1381
                                                                                                                                1561
                                                                                                                                    9357
                                                                                                                                         1621
                                                                                                                                              9297
                                                                                                                                                   1681
                                                                                                                                                        9237
                                                                                                                                                             1741
                                                                                                                                                                       1801
                                                                                                                                                                            9117
781
         841
                   901
                             961
                                       GTATCAGATACACACGGACTTCAACATTCTATCATAAGACCTACCCAACCCAACTGTTT
         gtal
                                  10016
                                                                                   9716
    10196
              10136
                                            9366
                                                      9686
                                                           1082
                                                                9836
                                                                     1142
                                                                         9116
                                                                              1202
                                                                                        1262
                                                                                             9656
                                                                                                  1322
                                                                                                       9836
                                                                                                                 9536
                                                                                                                      1442
                                                                                                                           9476
                                                                                                                                1502
                                                                                                                                    9416
                                                                                                                                         1562
                                                                                                                                              9326
                                                                                                                                                   1622
                                                                                                                                                        9536
                                                                                                                                                             1682
722
         782
                   842
                        10076
                             902
                                       962
                                                 1022
                                                                                                            1382
                                                                                                                                                                            9116
    QQ
              g
                        g
                                 QQ
                                           g
                                                    q
                                                          Q.Y
Db
                                                                        q
                                                                                  Ωp
                                                                                            g
                                                                                                Oy
Dp
                                                                                                           2 g
         ò
                   ò
                             Qγ
                                       à
                                               ò
                                                                   οχ
                                                                              ò
                                                                                       οŻ
                                                                                                                    ò
                                                                                                                         qq
                                                                                                                               oy
D
                                                                                                                                        ŏ
                                                                                                                                             QQ
                                                                                                                                                       g
                                                                                                                                                                 g
                                                                                                                                                                           g
                                                                                                                                                   å
                                                                                                                                                             ò
                                                                                                                                                                       ò
```

S

ò

g

셤

ò 셤 g ò 셤 ò q

à

```
Center: Whitehead Institute/ MIT Center for Genome Research Center: Ode: WIS the Wish share: WISR Web Site: http://www-seq.wi.mit.edu Contact: sequence_submissions@genome.wi.mit.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NOTE: This is a 'working draft' sequence. It currently consists of 14 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
                                                                                                                                                             Center project name: 14477

Center clone name: 24_13

Sequencing vector: M13; M77815; 100% of reads
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0,960731
Consensus quality: 156719 bases at least 040
Consensus quality: 166719 bases at least 030
Consensus quality: 163779 bases at least 020
Insert size: 170000; agarose-fp
Insert size: 170000; squrose-fp
                                                                                                                                                                                                                                                                                                                                                                                                                             Quality coverage: 4.4 in Q20 bases; sum-of-contigs
Quality coverage: 4.6 in Q20 bases; sum-of-contigs
    http://ftp.genome.washington.edu/RM/RepeatMasker.html
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               234 233: gap of 100 bp 100 bp 100 bp 2234 3987: contig of 1754 bp in length 3988 4087: gap of 100 bp 4088 4087: gap of 100 bp 4088 4087: gap of 100 bp 6972 7071: gap of 100 bp 10980 10979: contig of 2884 bp in length 10880 10979: gap of 100 bp 10980 15646: contig of 4667 bp in length 15647 15746: gap of 100 bp 100 bp 15747 25208: contig of 4667 bp in length 25209 25308: gap of 100 bp 25309 33458: contig of 8150 bp in length 46618 46717: contig of 13659 bp in length 46718 60565: contig of 13848 bp in length 46718 60565: contig of 13059 bp in length 6056 92690: contig of 13059 bp in length 25209 25309 app of 100 bp 60666 92690: contig of 13059 bp in length 46718 60565: contig of 13059 bp in length 122716: contig of 12925 bp in length 132716: contig of 12926 bp in length 132716: contig of 29256 bp in length 132716: contig of 29256 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           12271 122716: contig of 29920 up in 122717 122816: gap of 100 bp 122817 149492: contig of 26676 bp in length 149493 149592: gap of 100 bp 149593 165147: contig of 15555 bp in length. Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 867: contig of 867 bp in length
868 967: gap of 100 bp
968 2133: contig of 1166 bp in length
2134 2233; gap of 100 bp
2234 23387: contig of 1754 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /clone_lib="RPCI-11 Human Male BAC"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4088. .6971
/note="assembly_fragment"
7072. .10879
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note="assembly_fragment"
4088. .6971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note="assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-2413"

    .867
    /note="assembly_fragment

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           vector_side:left"
968. .2133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       clone_end:SP6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HTG 09-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACU21342 165147 bp DNA linear HTG 09-SEP-2000
Homo sapiens clone RP11-2413, WORKING DRAFT SEQUENCE, 14 unordered
pieces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 9, 2000 this sequence version replaced gi:6899755.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 165147)
                                                                                                                                                                                                                                                                                                                                                                                   8816 TTCAGGTGTTACTTGTGGATAAGCACAAATATTTCCTGTTTGGTTAAAACTTTAATCAGTT 8757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       caggtatccccccaaagaccccagaagtaaccctaggctcaatggagggtctagggacc 1981
1802 atggettttcaacatcacagecgaeceatatgagagggtggaeetatetaacaggtatee 1861
                                                                                                                  aggaatcgtgaagaagctcctacggaggctctcacagttcaacaaaactgcagtgccggt
                                                                                                                                                                                                                                     8996 CAGGTATCCCCCCAAAGACCCCAGAAGTAACCCTAGGCTCAATGGAGGGGTCTGGGGGACC
                                                                                                                                                                                                                                                                                       atggtatagagagagaaaccaagaaaaagaagccaagcaaaaatcaggctgagaaaaagca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ttcaggtgttacttgtggataagcacaaatatttcctgtttggttaaactttaatcagtt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                cttatctttcatctgtttcctaggtaaaccagcaaatttggctcgataatatcgctggcc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Nusbaum, C. and Lander, E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8696 TAAGCGTCAGGCTTGTTTTCATGCTGTGCCAC 8665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             taagcgtcaggcttgttttcatgctgtgccac 2252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Birren, B., Linton, L., Nusbaum
Homo sapiens, clone RP11-2413
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AC021342
AC021342.3 GI:10045513
HTG; HTGS_PHASE1; HTGS_DRAFT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (bases 1 to 165147)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                   2042
                                                                                                                                                                                                                                                                                                                                                                                                                                8876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AC021342/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                            9116
                                                                                            1862
                                                                                                                                      9026
                                                                                                                                                                                          1922
                                                                                                                                                                                                                                                                                    1982
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOURNAL
REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SOURCE
```

셤

a ò

ò

9

```
sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AC013692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 5
AC013692/c
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                               2101
                                 1442
                                                                  1502
                                                                                                   1562
                                                                                                                                    1622
                                                                                                                                                                                                                                        1802
                                                                                                                                                                                                                                                                                                                           100680
                                                                                                                                                                                                                                                                                                                                             1982
                                                                                                                                                                                                                                                                                                                                                                               2042
                                                                                                                                                                                                                                                                                                                                                                                                                                                 2161
1382
                101220
                                                 101160
                                                                                 101100
                                                                                                                                                     100980
                                                                                                                                                                      1682
                                                                                                                                                                                      100920
                                                                                                                                                                                                                                                                                                           1922
                                                                                                                                                                                                                                                                                                                                                              100620
                                                                                                                                                                                                                                                                                                                                                                                               100560
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  100380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VERSION
KEYWORDS
SOURCE
                                                 g
                                                                                  QΩ
                                                                                                                                                     g
                                                                                                                                                                                      qq
                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                              ga
                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                 ò
                                                                  ð
                                                                                                   ò
                                                                                                                                   ò
                                                                                                                                                                      ò
                                                                                                                                                                                                       δŏ
                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                Ω
                                                                                                                                                                                                                                 ï
                                                                                                                                                                                                                                                                                 Db 101880 GTATCAGATACACACCGGACTTCAACATTCTATCATAAGACCTACCCAACCCAACTGTT 101821
                                                                                                                                                                                                                                                                                                                                                                                      101101
                                                                                                                                                                                                                                                                                                                                                                                                                                                       TTATGGTTTCTATAACAACAGCATTATCATTTACTCTTCAGATAATGGTGGCCAGCCTAC 101281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ttatggtttctataacaacagcattatcatttactcttcagataatggtggccagcctac 1321
                                                                                                                                                                                                                                                                                                                                                                                                                       caatggcatatactccacacagatgtacactcagagagtacagcaaatcttagcttccca 1081
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            tgctgccatgcttcctgcttagatgaagcaatcaacaacgtgacattggctctaaagac 1261
                                                                                                                                                                                                                                                                                                                   901
                                                                                                                                                                                                                                                gtatcagatacacaccggacttcaacattctatcataagacctacccaacccaactgttt 781
                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                   ij
                                                                                                                                                                                                               Length 165147
                                                                                                                                                                               others
                                                                                                                                                                                                                                Indels
                                                                                                                                                                              1303
                                                                                                                                                                                                                                4
                                                                                                                                                                                                               5
                                                                                                                                                                                                               DB
                                                                                                                                                                              ų
                                             /note-"assembly_fragment"
3359. 46677
/note-"assembly_fragment"
/note-"assembly_fragment"
/note-"assembly_fragment"
/note-"assembly_fragment"
/note-"assembly_fragment"
/note-"assembly_fragment"
/123116. 149492
/note-"assembly_fragment"
/note-"assembly_fragment"
/note-"assembly_fragment"
/note-"assembly_fragment"
     10980. 15646
/note-"assembly_fragment"
15747. 25208
/note-"assembly_fragment"
25309. 33458
                                                                                                                                                                                                                               0; Mismatches
                                                                                                                                                                              51853
                                                                                                                                                                                                               Score 1207;
Pred. No. 0;
                                                                                                                                                   /note-"assembly_fragment
                                                                                                                                                                     vector_side:right"
29805 c 30056 g
                                                                                                                                                                                                              53.3%;
                                                                                                                                                              clone_end:T7
                                                                                                                                                                                                                       Best Local Similarity 99.7
Matches 1527; Conservative
                                                                                                                                                                              ಹ
                                                                                                                                                                              52130
                                                                                                                          misc_feature
      misc_feature
                        misc_feature
                                       misc_feature
                                                                         misc_feature
                                                                                          misc_feature
                                                                                                           misc_feature
                                                                                                                                           misc_feature
                                                         misc_feature
                                                                                                                                                                                                                Query Match
                                                                                                                                                                             BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1142
                                                                                                                                                                                                                                                 722
                                                                                                                                                                                                                                                                                 782
                                                                                                                                                                                                                                                                                                  101820
                                                                                                                                                                                                                                                                                                                   842
                                                                                                                                                                                                                                                                                                                                                                                      962
                                                                                                                                                                                                                                                                                                                                                                                                      101640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            101400
                                                                                                                                                                                                                                                                                                                                                                                                                       1022
                                                                                                                                                                                                                                                                                                                                                                                                                                       101580
                                                                                                                                                                                                                                                                                                                                                                                                                                                         1082
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        101520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           101460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             101340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               101280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1322
                                                                                                                                                                                                                                                                                                  요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                 ô
                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                      엄
                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
07-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostoml;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                     100741
                                           101161
                                                                                                                               ACTIGIGCACATCACTGACTGGTACCCCACTCTCATTTCACTGGCTGAAGGACAGATTGA 101101
                                                                                                                                                                                                                                                                                                                                                                                            100921
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1861
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACU13692 157043 bp DNA linear HTG 07-DEC HOMO sapiens clone RP11-21L8, *** SEQUENCING IN PROGRESS ***, unordered pieces.
100500 TTCAGGTGTTACTTGTGGATAAGCACAAATATTTCCTGTTTGGTTAAACTTTAATCAGTT
                                                                                                                                                                                                 101040 ACCCCGAGTAGATATTTGCATAACATTGACCCCATATACACCCAAAGGCAAAAAATGGCTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       100800 AIGGCTTTTCAACATCACAGCCGACCCATATGAGAGGGGGGGCCTACTAACAGGGTAGACATCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100740 AGGAATCGTGAAGAGCTCCTACGGAGGCTCTCACAGATCAACAAACTGCAGTGCCGGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CAGGTATCCCCCCAAAGACCCCAGAAGTAACCCTAGGCTCAATGGAGGGGTCTGGGGACC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAAGAAAAGCAAAAAAAAAGAAGAAGAACAGCAGAAAGCAGTCTCAGGTTCAACTTGCCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ttcaggtgttacttgtgggataagcacaaatatttcctgtttggttaaactttaatcagtt
                                                                                      acctgtgcacatcactgactggtaccccactctcatttcactggctgaaggacagattga
                                                                                                                                                                             tgaggacattcaactagatggctatgatatctgggagaccataagtgagggtcttcgctc
                                                                                                                                                                                                                                                                 accccgagtagatattttgcataacattgaccccatatacaccaaggcaaaaatggctc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1862 aggaatcgtgaagaagctcctacggaggctctcacaggttcaacaaaactgcagtgccggt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             caggtatcccccaaagaccccagaagtaaccctaggctcaatggaggggtctagggacc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      atggcttttcaacatcacagccgacccatatgagagggtggacctatctaacaggtatcc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a tggtatagagaggaaaaccaagaaaaagaagccaagcaaaaa tcaggctgagaaaaagca
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          taagcgtcaggcttgttttcatgctgtgccac 2252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AC013692.5 GI:11597077
HTG; HTGS_PHASE1.
```

```
97197: gap of 100 bp 103121: gap of 100 bp 103121: gap of 100 bp 108656: contig of 5535 bp in length 10876: gap of 100 bp 113420: gap of 100 bp 113420: gap of 100 bp 120045: gap of 100 bp 120045: gap of 100 bp 120045: gap of 100 bp 130194: contig of 9449 bp in length 130294: gap of 100 bp 130194: contig of 7562 bp in length 13786: contig of 7562 bp in length 13785: contig of 7562 bp in length 137956: gap of 100 bp 100 bp 100 bp 137956: gap of 100 bp 
                                                                                                                                                                                                                                                                                                                           p of 100 bp contig of 9659 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               p of 100 bp contig of 4062 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               100 bp -
5467 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       p of 100 bp contig of 1977 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                       in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BAC"
                                                                                                                                                                          2065
100 bp
.965 bp i
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /clone="RP11-21L8"
/clone_lib="RPCI-11 Human Male
                                                                                                           100 bp
f 2806 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100 bp
f 5702 bp i
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               100 bp
7741 bp i
                                                                                                                                                                                                                                                                                                                                                                                                            53557: gap of 100 bp
56770: contig of 3213 bp
56870: gap of 100 bp
59626: contig of 2756 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               of 100 bp
contig of 2521 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         p of 100 bp
contig of 3362 bp
                                                                                                                                                                                                                                                                                                                                                                    35: gap of 100 bp 53457: contig of 3122 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    of 100 bp
contig of 3205 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      p of 100 bp
contig of 4319 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                97: gap of 100 bp
87887: contig of 4390 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           oof 100 bp
contig of 3543 bp
                                                                                                                                                 30362: gap of 100 bp
32450: contig of 2088 bp
                                                                                                                                                                                                                                                                                 21: gap of 100 bp 40476: contig of 3055 bp
                      22898: gap of 100 bp
25224: contig of 2326 bp
                                                                p of 100 bp contig of 2045 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              vector_side:left"
1025. .1800
/note="assembly_fragment"
1901. .2800
/note="assembly_fragment"
                                                                                                                                                                                                                  contig of 1865
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note="assembly_fragment
clone_end:SP6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /organism="Homo sapiens"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               17225: gap of
154966: contig of
                                                                                                                                 contig of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                contig of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        contig of
                                                                                                                                                                                                                                                               : contig of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /db_xref="taxon:9606"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                              25225 25324: gap of
25325 27369: conti
27370 27469: gap of
27470 30262: conti
30263 30362: gap of
30363 32450: conti
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .v: gap of
97097
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    809: gap of
69014: cont
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       276: gap of 78978: cont
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             103022 103121: gap of
103122 108656: cont
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       gap of
                                                          324: gap of
27369: con+
                                                                                                                                                                                         32550: gap of
34415: con
                                                                                                                                                                                                                                                                                                                       40576: gap of
50235: con
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               59726: gap of
62247: con
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               69114: gap of 73176: con
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               78979 79078: gap of
79079 83397: con
                                                                                                                                                                                                                                      34515: gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            62347: gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           157043:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         91530:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  62709:
                                                                                                                                                                                                                                                             37321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  37957 137956:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     147226 154960
154967 155066:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       13321 113420:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               47126 147225:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91531 91630:
91631 9709
                                                                                                                                                                                                                                                                                                                                                                    50335:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            83497:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                113421 1206
120646 120745
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          30195 130294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62809:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         73276:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            08657 108756
                                                                                                                                                                                                                                                                               37421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       87987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97197
  21162
22799
22899
                                                                                                                                                                                                                                    34416
34516
37322
                                                                                                                                                                                                                                                                                                   37422
40477
40577
50236
                                                                                                                                                                                            32451
32551
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               59627
59727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            62248
62348
65710
65810
69015
69115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  83498
87888
87988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97098
                                                                                                                                                                                                                                                                                                                                                                                        50336
53458
53558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                83398
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           155067
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               56871
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                         Dipublished

Shades I to 157043)

Baldwin, J., Barnan, N. Beckerly, R. Boqualavkky, L., Boukhgalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Castle, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Castle, A., Castle, M., Collins, S., Collymore, A., Ferreitar, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehocky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Melfrim, J., Melfrim, J., Merlan, P., McGurk, A., Morman, C.H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Sutos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Tirreil, A., Vasalliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Lyopats, Were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Center: Whitehead Institute/ MIT Center for Genome Research
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NOTE: This is a 'working draft' sequence. It currently consists of 45 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Contact: sequence_submissions@genome.wi.mit.edu
------ Project Information
Center project name: L4053
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 15304 17901: gap of 100 bp 17902 19501: gap of 1659 bp in length 19561 19660: gap of 100 bp 19661 21061: contig of 1401 bp in length 21062 21161: qap of 100 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    of 100 bp
contig of 1620 bp in length
p of 100 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      100 bp
f 1129 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10238 10337: gap of 100 bp 101 bp 101338 12115: contig of 1778 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              of 100 bp contig of 1081 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100 bp
of 1599 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1446 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 15: gap of 100 bp
14480: contig of 2265 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        17801: contig of 1501 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100 bp
of 776 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100 bp
of 900 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           of 100 bp contig of 728 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100 bp
f 854 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  924: contig of 924 bp in length
                    Birren, B., Linton, L., Nusbaum, C. and Lander, E. Homo sapiens chromosome, clone RP11-21L8
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Web site: http://www-seq.wi.mit.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       contig of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            contig of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       contig of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          contig of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          contig of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  contig of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Center clone name: 21_L_8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12216 14480. 14481 14580: gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2900: gap of
3628: con
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3: gap of
6992: con
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2: gap of
8691: con
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οĮ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1801 1900: gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           gap of
  (bases 1 to 157043)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4682: gar
5763: '
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Center code: WIBR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3728: gap
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2800:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4582:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12215:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   16300:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5863:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7092:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8791:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2901
3629
3729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4583
4683
5764
5864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2116
               AUTHORS
TITLE
JOURNAL
REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                          TITLE
JOURNAL
REFERENCE
                                                                                                       AUTHORS
```

COMMENT

œ

```
Uppublished

State 1 to 165147)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F., Boukhgalter, B., Brown, A., Burkett, G., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArellano, K., Domino, M., Doyle, M., Fenestor, J., Ferreira, P., FitzHugh, M., Forrest, C., Gage, D., Galagan, J., Ferreira, P., FitzHugh, M., Forrest, C., Gage, D., Galagan, J., Ferreira, P., FitzHugh, M., Forrest, C., Gage, D., Galagan, J., Landers, T., Lehoczky, J., Levine, R., Ileu, C., Liu, G., Locke, K., Landers, T., Lehoczky, J., Levine, R., Ileu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., Merwan, P., McGurk, A., McKernan, K., Marquis, N., Meneus, L., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Olivar, T.M., Peterson, K., Pierre, N., Pisanl, C., Pollara, V., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Myman, D., Ye, W.J., Direct Subrission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AC021342 165147 bp DNA linear HTG 09-SEP-2000 HOMO saplens clone RP11-2413, WORKING DRAFT SEQUENCE, 14 unordered
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Sep 9, 2000 this sequence version replaced g1:6899755. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryogicus, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Bukamalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. I (bases 1 to 165147)
Birren, B., Linton, L., Nusbaum, C. and Lander, E. Homo sapiens, clone RP11-2413
                                             GGGGGCGCTGGCAGGATTCTGGATCCTCTGCCTCCTCACTTATGGTTACCTGTCCTGGGG 109850
                                                                                                                Db 109669 TGCCGAAGGAGTTAAACTGGAGAACTACTATGTCCAGCCTATTTGCACACCATCCAGGAG 109610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ------ Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
    gggggggctggcaggattctggatcctctgcctcctcacttatggttacctgtcctgggg 463
                                                                                                                                                                                                                                                                                                                                                         tgccgaaggagttaaactggagaactactatgtccagcctatttgcacaccatccaggag
                                                                                           ccaggccttagaagaagaagaagaagggccttactagctcaagctggagagaaactaga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Contact: sequence_submissions@genome.wi.mlt.edu
------ Project Information
Center project name: L4477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Center code: WIBR
Web site: http://www-seq.wi.mit.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HTG; HTGS_PHASE1; HTGS_DRAFT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                DD 109609 TCAGITIATIACTGGAAAGIA 109589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AC021342.3 GI:10045513
                                                                                                                                                                                                                                                                                                                                                                                                                                              tcagtttattactggaaagta 724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE
JOURNAL
    404
                                                                                         464
                                                                                                                                                                                                                                                                                                                                                                                                                                              704
                                             109909
                                                                                                                                                                              524
                                                                                                                                                                                                                       109789
                                                                                                                                                                                                                                                                                                                                                         644
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AC021342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                á
                                                                                         ò
                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Db 110269 TTTTTTTTTTTTCCTTGGTGGAAGCTGCTCTAGGGAGGAGGAGGAGGAGGAGGAGAAAAGTGAAA 110210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DD 110089 CAGGTGGAGAGTCTTCACCGTCTGTTGGGTGCATGTGTGCGCCCCCCCGCAGCGGCGCGGGCGC 110030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CGTGGTTCTCCGCGTGGAGTCTCACCTGGGACCTGAGTGAATGGCTCCCAGGGGCTGTGC 109970
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         caggiggagagicticaccgictgitigggigcatgigtgcgcccgccascggcgcgggggcg 283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                cgtggttctccgcgtggagtctcacctgggacctgagtgaatggctcccaggggctgtgc 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         44 ttttttttttccttggtggaagctgctctagggaaggggggaggaggaggaggagaaagtgaaa 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 157043;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 630; DB 2;
Pred. No. 0;
                                                                                                                                                                         7093. 8691
/note="assembly_fragment"
8792. 10237
/note="assembly_fragment"
1038. 12115
/note="assembly_fragment"
12216. 14480
/note="assembly_fragment"
14581 16200
/note="assembly_fragment"
16301. 17801
/note="assembly_fragment"
17902. 19560
/note="assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21162. .22798
/note="assembly_fragment"
22899. .25224
/note="assembly_fragment"
25325. .27369
/note="assembly_fragment"
27470. .30262
/note="assembly_fragment"
30363. .32450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /note-"assembly_fragment"
3251. 34415
/note-"assembly_fragment"
34516. 37321
/note-"assembly_fragment"
37422. 40476
/note-"assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Mismatches
                                                                                                                osto, .6992
/note="assembly_fragment"
1093, .8691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10577. .50235
/note="assembly_fragment"
                                                                                       /note="assembly_fragment"
5864, .6992
                    note="assembly_fragment"
                                                                note="assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note="assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  27.8%;
99.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 680; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    misc_feature
misc_feature
                                             misc_feature
                                                                                                                             misc_feature
                                                                                                                                                                              misc_feature
                                                                                                                                                                                                                       misc_feature
                                                                                                                                                                                                                                                                  misc_feature
                                                                                                                                                                                                                                                                                                              misc_feature
                                                                                                                                                                                                                                                                                                                                                         misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                              misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           misc_feature
                                                                                       misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Db 110209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Db 110029
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Db 109969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
```

/note="assembly\_fragment"

```
misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                         BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          40389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                            consists of 14 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
                                                                                                                                                                                                                                                                                                                                                                                                      NOTE: This is a 'working draft' sequence. It currently
                                                                                                                                                                                                                                              Insert size: 170000; agarose-fp
Insert size: 163847; sum-of-contigs
Quality coverage: 4.4 in Q20 bases; agarose-fp
Quality coverage: 4.6 in Q20 bases; sum-of-contigs
60566 60665: gap of 100 bp 60566 60665: gap of 100 bp 60666 92690: contig of 32025 bp in length 92691 92790: gap of 100 bp 92791 122717: contig of 29926 bp in length 122717 122816: gap of 100 bp 149493 149492: contig of 26676 bp in length 149493 149592: gap of 100 bp 149592: dap of 15555 bp in length.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       868 967: gap of 100 bp 968 2133: contig of 1166 bp in length 2134 2233: gap of 100 bp 2134 2233: gap of 100 bp 2134 2033: gap of 100 bp 971 200 bp 100 bp 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          46618 46717: gap of 100 bp 46718 60565: contig of 13848 bp in length 60566 60665: gap of 100 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              contig of 13059 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100 bp
9462 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25308: gap of 100 bp 33458: contig of 8150 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 867: contig of 867 bp in length
868 967: gap of 100 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Male
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              7234. .3987
/note="assembly_fragment"
4088. .6971
/note="assembly_fragment"
7072. .10879
/note="assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15747. .25208
/note="assembly_fragment"
25309 .33458
/note="assembly_fragment"
33559 .46617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /clone_lib="RPCI-11 Human
1. 867
/note="assembly_fragment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      note="assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        note="assembly_fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-2413"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        contig of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            vector_side:left"
968. .2133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15647 15746: gap of 15747 25208: cont
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              33558: gap of
46617: cont
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   clone_end:SP6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0980. .15646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25309
33459
33559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURES
```

```
HTG 09-FEB-2002
                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 39909 CAGGIGGAGAGTCTICACCGTCTGTTGGGTGCATGTGTGTGCGCCCCGCAGCGGCGGGGCG 39968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                39969 CGTGGTTCTCCGCGTGGAGTCTCACCTGGGACCTGAGTGAATGGCTCCCAGGGGCTGTGC 40028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                               164 ttctggaagattaaaagttgtcggacatggtgacagctgagaggaggaggatttcttgc
                                                                                                                                                                                                                                                                                                                                                                                                                                                   cgtggttctccgcgtggagtctcacctgggacctgagtgaatggctcccaggggctgtgc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   584 atttagagatgtgggttaccacggatctgagattaaaaacacctactcttgacaagctcgc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  644 tgccgaaggagttaaactggagaactactatgtccagcctatttgcacaccatccaggag
                                                                                                                                                                                                                                                                                                                                                                                                   caggtggagagtcttcaccgtctgttgggtgcatgtgtgcgcccgcascggcggggggg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ccaggccttagaagaaggaaggaagaaggggccttactagctcaagctggagaaaactaga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   524 gcccagcacaacttccacctcccagccccatctcattttcatcctagcggatgatcaggg
                                                                                                                                                                                                                           ö
                                                                                                                                                                                                 Length 165147;
                                                                                                                                                    1303 others
                                                                                                                                                                                                                           1; Indels
                                                                                                                                                                                                Score 630; DB 2;
Pred. No. 0;
0; Mismatches 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DNA
                                                                                                                                                    ų
175645 bp
                                                                                                                                      vector_side:right"
29805 c 30056 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                704 tcagtttattactggaaagta 724
                                                                                                                           clone_end:T7
                                                                                                                                                                                                  27.8%;
                                                                                                                                                                                                 Query Match 27.8
Best Local Similarity 99.9
Matches 680; Conservative
                                                                                                                                                    ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AC105414
                                                                                                                                                 52130
                           misc_feature
                                                  misc_feature
                                                                          misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AC105414/c
LOCUS
```

```
DEFINITION
ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                        404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                524
                  91286
                                                    104
                                                                                                                                                                                            224
                                                                                                                                                                                                                           91106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               90806
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AK027201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ρp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                    ò
                                                                                                                      ò
                                                                                                                                                        g
                                                                                                                                                                                            ò
                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                       Direct Submission
Submitted (04-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Feb 9, 2002 this sequence version replaced qi:18483567.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                  Homo saplens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Primates; Catarrhin1; Hominidae; Homo.
1 (bases 1 to 175645)
Waterston, R. H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  * NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N. but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
Homo sapiens chromosome 4 clone RP11-219L11, WORKING DRAFT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Chemistry: Dye-primer ET: 0% of reads
Chemistry: Dye-primer ET: 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 174714 bases at least 040
Consensus quality: 174899 bases at least 030
Consensus quality: 174899 bases at least 030
Insert size: 177000; agarose-fp
Insert size: 177000; agarose-fp
Ouality coverage: 10.91 in 020 bases; sum-of-contigs
Quality coverage: 8.85 in 020 bases; sum-of-contigs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 630; DB 2; Length 175645;
Pred. No. 0;
0; Mismatches 1; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                       Center: Washington University Genome Sequencing Center
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             74211: contig of 74211 bp in length 74311: gap of unknown length 175645: contig of 101334 bp in length.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequencing vector: M13; 0%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            109 others
                                                                                                                                                                                                                                                                                                                                                                                                    Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
                                                                                                                                                                                                                                                                                                                                                                     ----- Genome Center -----
                                                               HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          'note="assembly_name:Contig35"
32068 c 32415 g 56997 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note="assembly_name:Contig34"
74312. .175645
                                                                                                                                                                                         The sequence of Homo sapiens clone
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /organism="Homo sapiens"
/db_xref="taxon:9606"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /clone="RP11-219L11"
   SEQUENCE, 2 unordered pieces. AC105414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /chromosome="4'
                                                 AC105414.4 GI:18642912
                                                                                                                                                                                                                           (bases 1 to 175645)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               27.8%;
99.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 99.99
Matches 680; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1. .74211
                                                                                                                                                                                                                                            Waterston, R.H.
                                                                                                                                                                                                         Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   74312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 74212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54056
                                                                                       human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BASE COUNT
ORIGIN
DEFINITION
                                                                                                   ORGANISM
                                                                                                                                                      REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                TITLE
JOURNAL
                                ACCESSION
                                                 VERSION
KEYWORDS
                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                           TITLE
                                                                                                                                                                                                                                                                                                                                 COMMENT
                                                                                   SOURCE
```

ô

```
Action 1. Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Tanigami, A., Fujimara, Y., Nakajima, Y., Mizuno, T., Morinaga, M., Tanigami, A., Fujimara, T., Ono, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
NEDO human cDNA sequencing project Unpublished (2000)

2 (bases 1 to 1871)

Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T., Shibahara, T., Tanaka, T. and Nakamura, Y.
Shibahara, T., Tanaka, T. and Nakamura, Y.
Shibahara, T., Tanaka, T. and Nakamura, Y.
Submission
Submitted (29-AuG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRI 29-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    oligo capping; fis (full insert sequence).
Homo sapiens human lung cDNA to mRNA, clone_lib:LNG clone:LNG08487.
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GCCCAGCACAACTTCCACCTCCCAGCCCCATCTCATTTTCATCCTAGCGGATGATCAGGG 90747
                                                                                                                                                                                                                                    91107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     90867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     463
                                                                                                                                                                                                            91046 CGTGGTTCTCCGCGTGGAGTCTCACCTGGGACCTGAGTGAATGGCTCCCAGGGGTGTGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   90926 GGGGCCCTGGCAGGATTCTGGATCCTCTCCTCACTTATGGTTACCTGTCCTGGG
                                                                                                                                                                                                                                                                                                                                                                                        284 cgtggttctccgcgtggagtctcacctgggacctgagtgaatggctcccaggggctgtgc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gggggcgctggcaggattctggatcctctgcctcctcatggttacctgtcctgggg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                584 atttagagatgtgggttaccacggatctgagattaaaacacctactctgacaagctcgc
                                                                                                                                                                                 164 ttctggaagattaaaagttgtcggacatggtgacagctgagaggaggaggggtttcttgc
                                                                                                                                                                                                                                                                                       caggingagagicticaccgictgingtiggingtgingtgingconscances
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               geccageacaattecaceteccageeeeateteatttteateetageggatgateaggg
                                                                                ארסיבירטו 1871 bp mRNA linear
Homo sapiens cDNA: FLJ23548 fis, clone LNG08487.
AK027201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    90626 TCAGTTTATTACTGGAAAGTA 90606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       704 tcagtttattactggaaagta 724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AK027201.1 GI:10440271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 (sites)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE
JOURNAL
```

2212

2152

```
ACLU5383 176341 bp DNA linear HTG 26-JAN-2002
Homo sapiens chromosome 4 clone RP11-9G1, WORKING DRAFT SEQUENCE, 4
unordered pieces.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 176341)
Waterston, R.H.
           Direct Submission
Submitted (03-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St.
MO 63108, USA
On Jan 26, 2002 this sequence version replaced gi:18042331.
                                                          2093 acttgccattcaggtgttacttgtggataagcacaaatatttcctgtttggttaaacttt
                                                                                                                                                                     aatcagttcttatctttcatctgtttcctaggtaaaccagcaaatttggctcgataatat
                                                                                                                                                                                                                       744 AATCAGTTCTTATCTTTCATCTGTTTCCTAGGTAAACCAGCAAATTTGGCTCGATAATAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        * arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as as soon as it is available and the accession number will be preserved.
                                                                                    * NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequencing vector: plasmid; 66%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-primer ET; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 173694 bases at least 040
Consensus quality: 174243 bases at least 030
Consensus quality: 174824 bases at least 020
Insert size: 180000; agarose-fp
Insert size: 180000; agarose-fp
Unality coverage: 10.00 in 020 bases; agarose-fp
Quality coverage: 10.78 in 020 bases; sum-of-contigs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Center: Washington University Genome Sequencing Center
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .......... Genome Center
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1053: contig of 1053 bp in length
1153: gap of unknown length
5581: contig of 4428 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   cer code: WUGSC
site:http://genome.wustl.edu/gsc/index.shtml
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AC105383 AC031992
AC105383.2 GI:18376928
HTG; HTGS_PHASE1; HTGS_ACTIVEFIN.
                                                                                                                                                                                                                                                                            The sequence of Homo sapiens clone
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequencing vector: M13; 34%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (bases 1 to 176341)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Waterston, R.H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1054
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Center
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
                                                                                                                684
                                                                                                                                                                     2153
                                                                                                                                                                                                                                                                            2213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JOURNAL
REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                       AC105383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOURCE
                                                                                                                                                                                                                       Dp
                                                        ò
                                                                                                                q
                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                            ōλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /protein_de=BAB15689.1"
/db_xrefe="G1:10440212"
/db_xrefe="G1:10440212"
/db_xrefe="G1:10440212"
/db_xrefe="G1:10440212"
/db_xrefe="G1:10440212"
/db_xrefe="G1:10440212"
/db_xrefe="G1:10440212"
/db_xrefe="G1:1040212"
/db_xrefe="G1:1040212"
/db_xrefe="G1:10440212"
/db_xrefe="G1:104402"
/d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3
Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdnal@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Departent of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          aaaaagcaaaagaaaagc-aaaaaaagaagaagaaacagcagaaaagcagtctcaggttca 2092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   atggctcctgggcagcagctatgggatctggaacactgcaatccagtcagccatcagag 1674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gtgccggtcaggtatcccccaaaagaccccagaagtaaccctaggctcaatggaggggtc 1973
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            tagggaccatggtatagaggaggaaccaagaaaaagaagccaagcaaaatcaggctgag 2033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1495 agattgatgaggacattcaactagatggctatgatatctgggagaccataagtgagggtc 1554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ttcgctcaccccgagtagatattttgcataacattgaccccatatacaccaaggcaaaaa 1614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AGATTGATGAGGACATTCAACTAGATGGCTATGATATCTGGGAGACCATAAGTGAGGGTC 143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATGGCTCCTGGGCAGCAGCCTATGGGATCTGGAACACTGCAATCCAGTCAGCCATCAGAG 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GTGCCGGTCAGGTATCCCCCCCAAAGACCCCCAGAAGTAACCCTAGGCTCAATGGAGGGGTC 563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 263; DB 9; Length 1871;
Pred. No. 9.3e-143;
D; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                     /tissue_type="human lung"
/clone_lib="LNG"
204. -713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /note="unnamed protein product"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    546
                                                                                                                                                                                                                                                                                                                         /organism="Homo sapiens"
/db_xref="taxon:9606"
                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                 /clone="LNG08487"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11.6%;
99.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TCHSGVTCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 753; Conservative
                                                                                                                                                                                                                                                                                             .1871
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                Agency).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1854
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1914
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1974
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1675
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2034
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CDS
                                                                                                                                                                                                                                                                       FEATURES
                                                                                    COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

Louis,

```
misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                misc_feature
                                                                                                                                                                                                                                                                TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURES
                                                                                                                                                                                                                                                                                                                                 COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Diren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavkiy, L., Boukhgalter, B., Brown, A., Barna, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Illev, I., Johnson, R., Jones, C., Karatas, A., LaRocque, K., Landers, T., Lehoczky, J., Levine, R., Liu, G.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACO91322 232951 bp DNA linear HTG 11-JUN-2001
Mus musculus clone RP23-60D3, WORKING DRAFT SEQUENCE, 5 unordered
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. I. (bases 1 to 232951)
Birren, B., Linton, L., Nusbaum, C. and Lander, E. Mus musculus, clone RP23-60D3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2052 aaaaaaagaagaagaagaagcagcagaaagcagtctcaggttcaacttgccattcaggtgtta 2111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2172 tctgtttcctaggtaaaccagcaaatttggctcgataatatcgctggcctaagcgtcagg 2231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             97 AAAAAAAGAAGAAGAAACAGCAGAAAGCAGTCTCAGGTTCAACTTGCCATTCAAGTGTTA 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  217 TCTGTTTCCTAGGTAAACCAGCAAATTTGGCTCGATAATATCGCTGGCCTAAGCGTCAGG 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 176341;
5681: gap of unknown length
63724: contig of 58043 bp in length
53824: gap of unknown length
176341: contig of 112517 bp in length.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  302 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 99; DB 2; Length 176
Pred. No. 5.1e-46;
0; Mismatches 2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AC091322.2 GI:14336533
HTG; HTGS_PHASEI; HTGS_DRAFT; HTGS_FULLTOP.
                                                                                                                                                                                                                                                                                                                                                                                                             vector_side:left"

ture 63825. .176341

/note="assembly_name:Contig39"

61266 a 30790 c 29844 g 54139 t

    1053
    note="assembly_name:Contig33"
    1154. .5581

                                                                                                                                                                                                                                                                                  /note="assembly_name:Contig37
clone_end:SP6
                                                                                                                                                                                                                                                                                                                                                   5682. 63724
/note="assembly_name:Contig38
                                                                                                                                 /organism="Homo sapiens"
/db_xref="taxon:9606"
                                                                                       Location/Qualifiers
                                                                                                                                                                                               /clone="RP11-9G1"
                                                                                                                                                                                                                                                                                                                            vector_side:left"
                                                                                                                                                                              /chromosome="4"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2232 cttgttttcatgctgtgccac 2252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4.48;
99.08;
                                                                                                                                                                                                                                                                                                                                                                                                  clone_end:T7
                                                                                                             1. .176341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 99.0
Matches 199; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             house mouse.
Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Unpublished
  5582
5682
63725
63825
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AC091322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AC091322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            pieces.
                                                                                                                                                                                                                        misc_feature
                                                                                                                                                                                                                                                                misc_feature
                                                                                                                                                                                                                                                                                                                                                       misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                       misc_feature
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
AC091322/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
                                                                                       FEATURES
```

õ

g ð g à g ò g

```
McGrann, V., McGrann, K., McDeters, N., Melthews, C., McGrarny, M., McGrann, K., McGreters, N., Melthews, C., McGrarny, M., Minova, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oilver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Ribeack, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Sougnez, C., Spencer, B., Stanger-Thomann, N., Strauss, N., Subramanian, A., Talamas, J., Travers, M., Travis, N., Trigillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Lubmitted (15-ARR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 11, 2001 this sequence version replaced gi:13625486.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RW/RepeatMasker.html
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the places is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          154 2590; contig of 2437 bp in length 2591 2690; gap of 100 bp 2691 2690; gap of 100 bp 2691 11774; contig of 9084 bp in length 11775 11874; gap of 100 bp 11875 100524; contig of 8650 bp in length 100525 100524; contig of 8650 bp in length 100625 100524; contig of 132327 bp in length 1. 232951; contig of 132327 bp in length.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Contact: sequence_submissions@genome.wi.mit.edu
------- Project Information
Center project name: L13252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          53: contig of 53 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /clone_lib="RPCI-23 Female Mouse BAC"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Center code: WIBR
Web site: http://www-seg.wi.mit.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2691. 11774
/note="assembly_fragment"
11875. 100524
/note="assembly_fragment"
100625. 232951
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /note="assembly_fragment"
2691. .11774
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-60D3"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1. .53
/note="assembly_fragment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                vector_side:left"
154. 2590
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            clone_end:SP6
```

```
TITLE
JOURNAL
                                                                                                                             COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nature, D. M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alabbrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbaria, J., Babbrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbaria, J., Babbrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbaria, J., Babbrooks, S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbaria, J., Babbrooks, S., Bringel, R.L., Byrd, N.C., Carron, T.F., Burkett, C., Burrell, K.L., Byrd, N.C., Chen, R., Chen, Z., Chowdhy, I. Christopoulos, C. Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davida, M.L., Davis, C., Darbor, M.D., Dathorne, S.R., David, R., Davida, M.L., Davis, C., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C., Elhaj, C., Escotto, M., Ealls, T., Ferraguto, D., Flagq, M., Eddards, C., Harris, K., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Garoll, J.H., Gao, J., Garcia, A., Ganner, T., Garza, N., Gill, R., Garoll, J.H., Gaverar, W., Gunzarte, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Harris, C., Harris, K., Harris, C., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johuse, M., Marjue, F., Martin, R., Martindale, A., Louiseyed, H., Lozado, R., Homan, F., Martin, R., Martin, R., Martin, Man, C., Metzker, M., Man, M., Okwonuu, G., Oragunyen, A., Miller, G., Miller, R., Martin, R., Martin, R., Martin, M., Okwonuu, G., Oragunyen, R., Nickerson, E., Noser, M., Oragunyen, A., Tabor, P., Tang, H., Tanesy, J., Taylor, T., Taylor, T., Taylor, T., Tang, H., Tanesy, J., Taylor, T., Taylor, T., Warte, M., C., Wartington, S., Williamson M., Martington, S., Wi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCUSED 122992 bp DNA linear HTG 20-DEC-2001 Rattus norvegicus clone CH230-139G21, *** SEQUENCING IN PROGRESS ***, 50 unordered pieces.
                                                                                                                                                                                                                                                                         ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                           Db 146855 AAAATGGCTCCTGGGCAGCAGGCTATGGGATCTGGAACACTGCAATCCAGTC 146803
                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                          Oy 1611 aaaaatggctcctgggcagcagctatgggatctggaacactgcaatccagtc 1663
                                                                                                                                                                                                               Length 232951;
                                                                                                                                                  2.3%; Score 53; DB 2; Lengum -
ored, No. 7e-19;
ored 0; Indels
                                                                                           406 others
                                                            vector_side:right"
45474 c 46842 g 69434 t
                                                                                                                                                                                                                              100.0%; Prea. ...
/note="assembly_fragment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AC098604.4 GI:17973771
HTG; HTGS_PHASE1.
                                    clone_end:T7
                                                                                                                                                                                                                                       Best Local Similarity 100. Matches 53; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Norway rat.
Rattus norvegicus
                                                                                        70795 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AC098604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Rattus.
                                                                                                                                                                                                               Query Match
                                                                                        BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE
JOURNAL
REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AC098604
```

(bases 1 to 122992)

Worley, K.C.

AUTHORS

```
* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a "working draft' sequence. It currently

* Consists of 50 contigs. The true order of the pieces

* is not known and thair order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.
                                                                                                                                                                                           findPhrapList
Consensus quality: 115001 bases at least Q40
Consensus quality: 121003 bases at least Q30
Consensus quality: 125541 bases at least Q20
Extimated insert size: 103320; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 1.9x in Q20 bases; sum-of-contigs estimation
Submitted (26-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, Tx 77030, USA on Dec 20, 2001 this sequence version replaced gi:17064604.
                                                                                                                                                                                   S
                                                                                                                                                                    ------ Summary Statistics
Assembly program: Phrap; version 0.990329First call
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18920: contig of 5998 bp in length 19020: gap of unknown length 19020: gap of unknown length 18771.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                              length
bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  bp in length
length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       length
bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   unknown length
of 2102 bp in length
unknown length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ength.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ength.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ength
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ength
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ength
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ength
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         length
                                                                 Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          pp in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        pp in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                do in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               of 4721 b
of 4721 b
unknown l
                                                                                                                                                                                                                                                                                                                                                                                                                                                              unknown of 7434 h
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     unknown
of 5251 h
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     unknown
of 4648 h
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 unknown
of 1920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               unknown of 1672 h
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              unknown
of 3138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            unknown
of 1622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       unknown
of 1356
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          unknown
of 2686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    unknown
of 1727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          unknown
of 3184
                                                                                                                                                                                                                                                                                                                                                                                                                                                 5288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  unknown
                                                                                                                                       Center project name: GIKS
Center clone name: CH230-139G21
                                                                                                                                                                                                                                                                                                                                                                                                                                              5288: contig of 5
5388: gap of unkn
12822: contig of 7
12922: gap of unkn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               oţ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 contig (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          contig
gap of u
                                                        ---- Genome Center
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  contig
qap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24271: contig
24371: gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             29092:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                41472:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                55045:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               70827:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     44856:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5389
112823
112823
118921
119021
24272
24372
224372
234373
33421
33421
33421
338335
441573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 44857
49505
49605
51168
511868
511868
511868
511868
511868
511868
511868
511868
511868
511868
511868
511868
511868
511868
511868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  63416
63516
65243
65343
67015
67115
68526
68626
```

```
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK plate: a Column: 7 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF
                                                                                                                                                                                                                                                                                                                                                                                                              cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web Site:
Dittp://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /translation="DASAHASAHASVDIPGPHAAWQISEKWFNSCSGAVCLSIPSGSL
PLTSKGSLSCLRLVSYMGL"
189 c 189 g 282 t
                                                                                             PRI 22-OCT-2001
                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostom1;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 927)
                                                                                                                                                                                                                                                                     Direct Submission
Submitted (29-J0N-2001) National Institutes of Health, Mammallan
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /product-"Unknown (protein for IMAGE:3890809)"
/protein_id="AAH09528.1"
/db_xref-"GI:16306921"
                                                                                            BC009528 927 bp mRNA linear P Homo sapiens, clone IMAGE:3890809, mRNA, partial cds. BC009528.1 GI:16306920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /db_xref="taxon:9606"
/clone="INAGE:3890809"
/tissue_type="Lung, carcinoma, large cell
undliferentiated."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 9; Length 927, 0.0069; Indels
                                                                                                                                                                                                                                                                                                                                                  NIH-MGC Project URL: http://mgc.nci.nih.gov
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /note="Vector: pCMV-SPORT6"
<1. .194</pre>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1.1%; Score 26; DB 100.0%; Pred. No. 0.0
                                                                                                                                                                                                                                                                                                                                                                                Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
 Db 119477 GAGAACTACTATGTCCAGCCTATTTGCAC 119505
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /organism="Homo sapiens"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /clone_lib="NIH_MGC_69"
/lab_host="DH10B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Location/Qualifiers
1. .927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                <1. .194
/codon_start=3</pre>
                                                                                                                                                                                                                                                                                                                                                                    Contact: MGC help desk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 1.1
Best Local Similarity 100.
Matches 26; Conservative

    (bases 1 t
    Strausberg, R.

                                                                                                                                                                                             Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        analysis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R. M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13
                                                                                                          DEFINITION
                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                       TITLE
JOURNAL
                                                                                                                              ACCESSION
                                                                                                                                          VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                         AUTHORS
                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 1
BC012659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CDS
                                                                              BC009528
                                                                                                                                                                                                                                                                                                                                                    REMARK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                   COMMENT
                                                                RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 122992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4990 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 29; DB 2; Length 122
Pred. No. 0.0001;
0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        in length.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           length
                                                                                                                            in length
                                                                                                                                                                                        bp in length
                                                                                                                                                                                                                                                                                                                                                                              bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        bp in length
                                                                bp in length
                                                                                               bp in length
                                                                                                                                                          bp in length
                                                                                                                                                                                                                         in length
                                                                                                                                                                                                                                                        bp in length
                                                                                                                                                                                                                                                                                       length
                                                                                                                                                                                                                                                                                                                     bp in length
                                                                                                                                                                                                                                                                                                                                                    bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                              length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      length
bp in le
                                                                                                                                                                                                                                                                                                   length
                                                                                                                                                                                                                                                                                                                                                                                                                              length
                                               length
                                                                                                                                                                                                                                                                     ength.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           length
                                                                              length
                                                                                                              length
                                                                                                                                             length
                                                                                                                                                                            ength
                                                                                                                                                                                                        length
                                                                                                                                                                                                                                       length
                                                                                                                                                                                                                                                                                                                                     length
                                                                                                                                                                                                                                                                                                                                                                 ength
                                                                                                                                                                                                                                                                                                                                                                                                length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ength
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ļ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Į.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               pp in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           pp in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                pp tu
                                                                                                                                                                                                                                                                                         ņ
                                                                                                                                                                                                                                                                                                                                                                                                                                                pb in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Į,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /organism="Rattus norvegicus"/db_xref="taxon:10116"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /clone="CH230-139G21"
23990 c 23255 g 35545 t
                                                                of 1841 h
unknown J
of 1865 h
                                                                                                                                                                                                                                       unknown
of 2236
                                                                                                                                                                                                                                                                                                                                    unknown
of 1725
                                                                                                                                                                                                                                                                                                                                                                 unknown of 1615 hunknown of 1589 h
                                                                                                                                                                                                                                                                                                                                                                                                                            unknown
of 1125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   unknown
of 1418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                unknown
of 1397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         unknown
of 1276 l
                                                                                                                                                                                                                                                                       unknown
of 1699
                                                                                                                                                                                                                                                                                                      unknown
of 1690
                                                                                                                                                                                                                                                                                                                                                                                                                                                             unknown
of 1837
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         unknown
of 2287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       unknown
of 1542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1600
                                                                                                                           1210
                                                                                                                                                                                                        unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         of 1115
                                                                                                            unknown
                                                                                                                                                          of 1728
                                                                                                                                                                            unknown
                                                                                                                                                                                         1535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          of 1504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   of 1120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   of 1504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 of 1403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             of 1186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        of 1281
                                                 unknown
                                                                                                                                                                                             oţ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          oţ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οĘ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
                                              gap of
contig
gap of
                                                                                            contig
gap of
                                                                                                                                                                                                                                                                                                                 contig
gap of
contig
                                                                                                                                                                                                                                                                                                                                                                                contig
gap of
contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       contig
gap of
contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             gap of
contig
                gap of
contig
                                                                                                                                                          contig
gap of
                                                                                                                                                                                                                                                                       gap of
contig
                                                                                                                                                                                                                                                                                                                                                                                                                            gap of
contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             gap of
contig
                                                                                                                                                                                                                                                                                                    gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                            gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        contig
                                                                                                                           contig
                                                                                                                                                                                         contig
                                                                                                                                                                                                                                                        contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gap of
                                                                                                                                                                                                                         contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           gap of
                                                                                                                                           gap of
                                                                                                                                                                                                                                       gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gap of
                                                                                                                                                                                                           gap of
                                                                                                                                                                                                                                                                                                                                                                   gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1.3%; Scor.
100.0%; Pre
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     107150:
108692:
                                                                             78783:
80648:
80748:
                                                                                                                                                                                                                                                                                                                                                 95065:
95165:
96780:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            119120:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21616:
                                              76842:
78683:
                                                                                                                           81958:
82058:
83786:
83886:
                                                                                                                                                                                                                                       87415:
89651:
                                                                                                                                                                                                                                                                     89751:
91450:
                                                                                                                                                                                                                                                                                                                93240:
                                                                                                                                                                                                                                                                                                                                                                                                                              98269:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .08792:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     109912:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     110012:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             117834:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             119020:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        122992:
                  3851:
                                                                                                                                                                                                                                                                                                    91550:
                                                                                                                                                                                                                                                                                                                                                                                               96880:
                                                                                                                                                                                                                                                                                                                                                                                                                                              99694:
                                                                                                                                                                                                                                                                                                                                                                                                                                                             : 46166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        103235:
                                                                                                                                                                                                                                                                                                                                                                                                                 98469:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               116134
                                                                                                                                                                                         85421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            101631
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        104763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
Matches 29; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               110013
111517
111617
113035
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               113135
114538
114638
116035
116135
117735
117835
70828
73752
73852
76743
76843
78684
90649
80649
81959
82059
83787
83887
                                                                                                                                                                                                       85422
87522
87522
87522
87522
89452
89752
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
91451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  108793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         119121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        120336
121617
121717
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    35212 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURES
```

ö

663 gagaactactatgtccagcctatttgcac 691

ð

```
Buday, D. M., Adams, C., Adio-Oduola, B., Ali-osman, F. R., Allen, C., Alstrooks, S. L., Amaratunge, H. C., Are, J. R., Ayele, M., Banks, T., Barbaria, J. Benton, J. Binage, K., Blankenburg, K., Bonnin, D., Budack, C., Burch, P., Britan, M., Bryant, N. B., Banken, J. Barbaria, J. Benton, J. Binage, K., Blankenburg, K., Bonnin, D., Bouck, J., Britan, C., Burch, R., Bryant, N. C., Carron, T. E., Carter, M., Cavazos, S. R., Chacko, J., Changer, R., David, M. D., Dathorne, S. R., David, R., Davila, M. L., David, M. D., Dathorne, S. R., David, R., Davila, M. D., Dathorne, S. R., David, R., Davila, M. L., Davis, C., Coyle, M. D., Dathorne, S. R., David, R., Davila, M. L., Dray, C., Elbaj, C., Escotto, M., Earnhart, C., Edgar, D., Edgar, M., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Garlis, T., Gavaza, M., Fallas, T., Hantlon, K. J., Harris, K., Harris, K., Hauber, J., Haylak, P., Hawes, A., Hennadez, J., Harris, K., Harris, M., Holues, M., Holluns, B., Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Martinez, E., Marsoy, E., Martin, R., Martin, R., Martin, R., Ma, J., Liz, L., Lichtarge, O., Lieu, C., Liu, J., Liu, R., Martinez, E., Marsoy, E., Martin, R., Martin, R., Martin, R., Mayla, M., Okwunu, G., Martin, R., Martin, R., Martin, R., Mayla, M., Okwunu, G., Martin, R., Marti
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Submitted (27-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, Tx 77030, USA
                                                                                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Rattus norvegicus clone CH230-109H4, *** SEQUENCING IN PROGRESS ***, 26 unordered pieces.
                                                                                                                                                                                                                                              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequencing vector: Plasmid; M77789
Chemistry: Dye-terminator Big Dye: 98% of reads
Assembly program: Phrap; version 0.990329First call to findPhrapList
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Consensus quality: 14825 bases at least Q40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Center: Baylor College of Medicine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Center project name: GPMZ
                                                                                              AC108339.1 GI:18377119
                                                                                                                                                                                                                                                                                                         (bases 1 to 28735)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (bases 1 to 28735)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Center code: BCM
                                                                                                                                                   Norway rat.
Rattus norvegicus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Direct Submission
                                                                                                                             HTG; HTGS_PHASE1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Worley, K.C.
         DEFINITION
                                                                                                                                                                                     ORGANISM
                                                                 ACCESSION
                                                                                          VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                             REFERENCE
                                                                                                                                                                                                                                                                                                                                         AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: http://image.llnl.gov Series: IRAK Plate: 18 Row: f Column: 6 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 3309175. Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ELELENAVQELAEGANAAYLOLLNIFAYGTYPDYTANKESLPELSAAQONKLKHLTIV
SLASRMKCIPYSVLKDLEMRNIRELEDLITEAVYTDIIQGKLDQRNDLLEVDFCIGR
DIKKDINIVVYTHEWGGGEAVLIGIEGOVLRANQYKENHHRTQQOVEAEVSNIKK
TLKATASSSAQEMEGLERECPPHTEQRQPTKKMSKVKGLVSSRH"
541 c 516 g 481 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /product-"Similar to COP9 (constitutive photomorphogenic), subunit 7 P. (Arabidopsis)" /protein_id="ARABI2659.1" /db_xref="GI:15215085"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /translation="MAGEQKPSSNLLEQFILLAKGTSGSALTTLISQVLEAPGVYVFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HTG 27-JAN-2002
      ROD 20-AUG-2001
                                                                                                                                                                                                                                        Mus musculus
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2003)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Contact: MGC help desk
Email: cgapbs-rémail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M. A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
Contact: amadanésystemsbiology.org
contact: amadanésystemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Ketteman and Anuradha Madan
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /organism="Mus musculus"
/organism="Mus musculus"
/db xref="taxon:10090"
/clone="Mc:13817 IMAGE:4017012"
/tissue_type="Mammary tumor metastatized to lung. Tumor arose spontaneously from a senescent normal mammary cloneallo outgrowth infected with the virus MMTV."
/clone_lib="NCI_CGAP_Lu29"
/lab_host="DH10B"
                                                                                                                                                                                                                                                                                                                                                                                                                    Submitted (15-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
BC012659 2003 bp mRNA linear ROD 20-AUG Mus musculus, Similar to COP9 (constitutive photomorphogenic), subunit 7b (Arabidopsis), clone MGC:13817 IMAGE:4017012, mRNA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1.1%; Score 26; DB 10; Length 2003; 100.0%; Pred. No. 0.0067; Live 0; Mismatches 0; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NIH-MGC Project URL: http://mgc.nci.nih.gov
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /note="Vector: pCMV-SPORT6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28735 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BC012659.1 GI:15215084
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 1.1%
Best Local Similarity 100.0
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .911
                                                                                                                                                                                                                                                                                                                                                                                              Direct Submission
                                                                                          complete cds.
                                                                                                                                                                                                                                                                                                                                                               Strausberg, R
                                                                                                                                                                                                              house mouse.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AC108339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 14
AC108339/c
LOCUS
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
ORIGIN
                                                                                                                                                                                                                                           ORGANISM
                                                                                                                                                                                                                                                                                                                                                            AUTHORS
TITLE
                                                                                                                    ACCESSION
                                                                                                                                                VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                       JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REMARK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
```

INV 04-NOV-2000

ö

Gaps

ö

```
Direct Submission
Submitted (04-NOV-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
Submitted by:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

Location/Qualifiers
1. .38866
                                                                                                                                                                                                                                                                                          CBRG36E21 38866 bp DNA linear INV 0 Caenorhabditis briggsae cosmid G36E21, complete sequence. AC084563.1 GI:11095013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Genome Sequencing Center
Department of Genetics, Washington University,
T. Louis, MO 63110, USA
e-mail: jspieth@watson.wustl.edu
                                                                                                                                                                                                                                                                                                                                                                                         Caenorhabditis briggsae.
Caenorhabditis briggsae
Eukaryota; Metazoa: Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1.1%; Score 25; DB 3; Length 38866; 100.0%; Pred. No. 0.024; tive 0; Mismatches 0; Indels (
                                                                   Length 28735;
                                                                                                           0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 (bases 1 to 38866)
Washington University Genome Sequencing Center.
The C. briggsae Genome Sequencing Project
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /organism="Caenorhabditis briggsae"
/strain="GujArat G16"
/db_xref="Laxon:6238"
                                                                   DB 2;
                                                                 Score 25; DB 2; Pred. No. 0.02
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7483 g 11766 t
                                                                                                                                               Search completed: July 23, 2002, 19:09:16
Job time: 14232 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9383 AAGAAAAGCAAAAAAAAAAAAGAAGAAGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2043 aagaaaagcaaaaaaaaagaagaa 2067
                                                    1.1%; Scc.
100.0%; Pre
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    One="G36E21
6629 c 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2 (bases 1 to 38866)
Waterston, R.
                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
                                                                                      Best Local Similarity
Matches 25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         12988 a
                                                                   Query Match
                                                                                                                                                                                                                                                                         CBRG36E21/c
                                                                                                                                                                                                                                                                                                               DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOURNAL
REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITEE
JOURNAL
                                                                                                                                                                                                                                                                                                                                  ACCESSION
                                                                                                                                                                                                                                                                                                                                                     VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                     NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

NOTE: This is a "working draft" sequence. It currently consists of 26 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
Consensus quality: 18227 bases at least Q30 Consensus quality: 20289 bases at least Q20 Estimated insert size: 9365; sum-of-contigs estimation Quality coverage: Ox in Q20 bases; agarose-fp estimation Quality coverage: 0.1x in Q20 bases; sum-of-contigs estimation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           unknown length
of 1028 bp in length
unknown length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               unknown length
of 1163 bp in length
unknown length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        unknown length
of 1124 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     unknown length
of 1177 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     of 1756 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    of 1058 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             of 1568 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                of 1678 bp in length
                                                                                                                                                                                                                                                                                                                                                     unknown length
of 862 bp in length
unknown length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      of 818 bp in length
unknown length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            of 690 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                of 771 bp in length
                                                                                                                                                                                                                                                                                                                                    contig of 884 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                              of 856 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                      of 783 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              979 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    of 821 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        of 874 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          of 908 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            of 828 bp in length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 of 1323 bp in l
unknown length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          unknown length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           unknown length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     unknown length
                                                                                                                                                                                                                                                                                                                                                                                                                                     length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           unknown length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   unknown length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              914 bp in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      882 bp in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              of 409 bp in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            of 457 bp in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-10944"
                                                                                                                                                                                                                                                                                                                                                                                                                                   unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              of
                                                                                                                                                                                                                                                                                                                                                     gap of contig co
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          contig
gap of
contig
gap of
contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gap of
contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             contig
gap of
gap of
contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   contig
gap of
contig
gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gap of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                contig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             gap of
                                                                                                                                                                                                                                                                                                                                9884:
1946:
2802:
2802:
3785:
3785:
4864:
6936:
77036:
77036:
8018:
8475:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13115:
13933:
14033:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21183:
22939:
23039:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25233:
25333:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11498:
11598:
12506:
12606:
13015:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                23810:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                23910:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                27011:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            14861:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 14961:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19515:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1083:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9497
10371
10471
11499
11599
12507
12607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16915
17015
18139
18239
19416
19516
21084
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21184
22940
23040
                                                                                                                                                                                                                                                                                                                                                                                         11847
229803
336863
33686
33686
47686
6023
6023
7037
7037
7037
7037
8819
8810
9397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4034
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FEATURES
```

ö

Gaps

ö

2573 others

7614 t

5034 g

4798 c

8716 a

BASE COUNT

			1. 14	* 1
	•			
			•	
	•			
*)				
		4		
	.°			
		A		
		•		
-				
	·			
- X - *	•			
•				
•				
		· 3		
•	i i	*		

```
Homo sapiens
                                                                                                                                                                                                        agaaaaagaagccaagcaaaaatcaggctgagaaaaggcaaaagaaaagc-aaaaaaga 2060
                                                                                                                                                                                                                2568 agaaaaagaagccaagcaaaatcaggctgagaaaagcaaaagcaaaaagaaaaga 2627.
                                                                                                                                                                        ccagaagtaaccctaggctcaatggagggtctagggaccatggtatagagaggaacca 2001
                                                                                                                                                                                                2508 ccagaagtaaccctaggctcaatggagggtctggggaccatggtataaagaggaaacca 2567
                                                                                                                                        ccgacccatatgagagggtggacctatctaacaggtatccaggaatcgtgaagaagctcc 1881
                                                                                                                                                                 2388 ccgacccatatgagagggtggacctatctaacaggtatccaggaatcgtgaagaagctcc 2447
                                                                                                                        1522
                   1282
                                                                                                                                                                                                                                         a
                                                                                                                                                                                                         g
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                        ద
                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                  õ
                                                                                                                                                                                        셤
                                                                                                                                                                         g
                                                                                                                                                                                  ò
                                                                                                                                                          g
                                                                                                                                                                  ò
                                                                                                                                         g
                                                                                                                                                   ö
                                                                                                                           8
                                                                                                                                   ò
                                                                                                          ă
                                                                                                                   ≈
                                                                                          Q
                                                                                                    <u>≻</u>
```

Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation; probe; blood coagulation disorder; cancer; cellular adhesion disorder; secreted protein; transmembrane protein; ss.

Human PRO708 nucleotide sequence.

(first entry)

07-DEC-1999 AAZ33987;

BP

AAZ33987 standard; cDNA; 4650

4

AAZ33987 RESULT

```
98US-0083559.
98US-0083742.
98US-0084366.
                                                                                                                                                                                                                                                                                                                                                                                                                                          98US-0084414.
98US-0084441.
98US-0084598.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98US-0084627.
98US-0084637.
                                                                                                                                                                                                                                                                                                                                                                 98US-0083392.
98US-0083495.
98US-0083496.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0084600
                                                                                                                                                                                                                                                                                                                                                                                                   980S-0083545
                                                                                                                                                                                                                                                                                                                                                                                                                980S-0083558
                                                                                                                                                                                                                                                                                                                                                                                      98US-0083499
98US-0083500
                                                                                                                                                                                                                                                                                                                                                           98US-0083322
                                                                                                                                                                                                                                                                                                                                                      98US-0083336
                                                                                                                                                                                                                                                                                                                     98US-0082700
                                                                                                                                                                                                                                                                                                                                              98US-0082796
                                                                                                                                                                                                                                                                                     98US-0081838
                                                                                                                                                                                                                                                                                                 98US-0081955
                                                                                                                                                                                                                                                                                                               98US-0082569
                                                                                                                                                                                                                                                                                                                                  98US-0082804
                                                                                                                                                                                                                                                                                                                                         980S-0082767
                                                                                                                                                                                                                                                                                                        98US-0082568
                                                                                                                                                                                                                                                                                                                            98US-0082704
                                                                                                                                                                                                                                                                                            98US-0081952
                                                                                                                                                                                                                                                                        98US-0081229
                                                                                                                                                                                                                                        98US-0081049
                                                                                                                                                                                                                                                          98US-0081195
                                                                                                                                                                                                                                                                98US-0081203
                                                                                                                                                                                                                                                                                98US-0081817
                                                                                                                                                                                                                                                    98US-0081071
                                                                                                                                                                                                                         98US-0080333
                                                                                                                                                                                                                                   98<del>0S-</del>0080334
                                                                                                                                                                                    9803-0080105
                                                                                                                                                                                                 98US-0080165
                                                                                                                                                                                                        98US-0080194
                                                                                                                                                                                                               98US-0080327
                                                                                                                                                    980S-0079689
                                                                                                                                                                                           98US-0080107
                                                                                                                                                           98US-0079728
                                                                                                                                                                98US-0079786
                                                                                                            980S-0078936
                                                                                   98US-0078004
98US-0040220
                                                                                                       98US-0078910
                                                                                                                    98US-0078939
                                                                                                                                              980S-0079664
                                                                                                  980S-0078886
                                                                                                                                        980S-007966
                                                   98US-0077450
                                                                        98US-0077649
                                                                 98US-0077641
                                                                               98US-0077791
                                       99WO-US05028
                                                                                                                                                                               )0-SD86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      07-MAY-1998;
07-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                            06-MAY-1998
                                                                                                                                                                                                                                                                                                                                                                                               29-APR-1998
                                                                                                                                                                                                                                                                                                                                                                                                      APR-1998
                                                                                                                                                                                                                                                                                                                                                                                                            39-APR-1998
                                                                                                                                                                                                                                                                                                                                                                                   29-APR-1998
                                                                                                                                                                                                                                                                                                                                                  23-APR-1998
                                                                                                                                                                                                                                                                                                                                                        27-APR-1998,
                                                                                                                                                                                                                                                                                                                                                               28-APR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                30-APR-199
                                                                                                                                                                                                                                                                                                                                                                                                                                     05-MAY-199
                                                                                                                                                                                                                                                                                                                               22-APR-1998
                                                                                                                                                                                                                                                                                                                                     22-APR-1998
                                                                                                                                                                                                                                                                                                                                            23-APR-1998
                                                                                                                                                                                                                                                                                                      15-APR-1998
                                                                                                                                                                                                                                                                                                           21-APR-1998
                                                                                                                                                                                                                                                                                                                   21-APR-1998
                                                                                                                                                                                                                                                                                                                       22-APR-1998
                                                                                                                                                                                                                                                                                  15-APR-1998
                                                                                                                                                                                                                                                                                        15-APR-1998,
                                                                                                                                                                                                                                                                           39-APR-1998
                                                                                                                                                                                                                                                 38-APR-1998,
                                                                                                                                                                                                                                                        08-APR-1998
                                                                                                                                                                                                                                           08-APR-1998
                                                                                                                                                                                                           31-MAR-1998
                                                                                                                                                       -MAR-1998
                                                                                                                                                              -MAR-1998
                                                                                                                20-MAR-1998;
20-MAR-1998;
                                                                                                                                                                    27-MAR-1998
                                                                                                     20-MAR-1998
                                                                                                           20-MAR-1998
                 WO9946281-A2
                                                                                        13-MAR-1998
                                                                                              17-MAR-1998
                                          08-MAR-1999;
                                                                           11-MAR-1998
                                                                                 L2-MAR-1998
                                                                    11-MAR-1998
                              16-SEP-1999
                                                       10-MAR-1998
```

: फेरब्रुड -	* t '' *	A Company of the Comp	
	·		
	ar · · · · ·		
;			
		*	
		* •-	
# <sup>3</sup>	•		
* **:			
			• •
	. (1)		
,			
			·
			. **
			•
		, and the second se	
	÷		*
y.	**		
1			
2.5			
(X)			
*			
•	*		
	: *		<b>\</b>
			•
	*		
A. The second			

```
New secreted and transmembrane polypeptides and their polynucleotides, useful for treating blood coagulation disorders, cancers and cellular adhesion disorders -
                                                                                                                                                                                                             Chen
                                                                                                                                                                                                             KΡ,
                                                                                                                                                                                                                                                                                                                                                                                            G; 1235 T;
                                                                                                                                                                                                              Baker
                                                                                                                                                                                                             'n
                                                                                                                                                                                                                                                                                                                                                                                            C; 1026
                                                                                                                                                                                                             Yuan
                                                                                                                                                                                                                                                                                 2; Fig 42; 530pp; English
                                                                                                                                                                                                              À
                                                                                                                                                                                                                                                                                                                                                                                            Sequence 4650 BP; 1431 A; 958
                                                                   980S-0085582
980S-0085689
980S-0085697
980S-0085700.
980S-0085704.
980S-0086123.
980S-0086414.
980S-0086416.
                                                                                                                                                        98US-0087106.
98US-0087208.
98US-0094651.
                                                                                                                                                                                                              Gurney
98US - 0084639.
98US - 0084643.
98US - 0084643.
98US - 0085323.
98US - 0085338.
98US - 0085339.
                                                              980S-0085580
                                                                                                                                                                              98US-0100038
                                                                                                                                                                                               (GETH ) GENENTECH INC
                                                                                                                                                                                                             Wood WI, Goddard A,
                                                                                                                                                                                                                            WPI; 1999-551358/46.
P-PSDB; AAY41701.
              07 - MAY - 1998;
13 - MAY - 1998;
13 - MAY - 1998;
13 - MAY - 1998;
15 - MAY - 1998;
                                                                                                                                 22-MAY-1998;
22-MAY-1998;
28-MAY-1998;
                                                                                                                                                        28-MAY-1998;
28-MAY-1998;
30-JUL-1998;
11-SEP-1998;
                                                                   -MAY-1998;
-MAY-1998;
-MAY-1998;
-MAY-1998;
                                                                                                  -MAY-1998;
-MAY-1998;
-MAY-1998;
-MAY-1998;
                                                              15-MAY-1998
                                                                                                                                                                                                                                                                                  Claim
```

'n

```
ä
The present invention describes secreted and transmembrane polypeptides and their polynucleotides. The nucleotide sequences are useful as sources of probes, primers, for chromosome mapping, and for generation of antisense sequences. They can also be used to create transgenic animals. The proteins can be used to treat a variety of diseases and disorders, depending on their function. Diseases that may be treated include abload coagulation disorders, cancers and cellular adhesion disorders. They may also be used to raise antibodies. AA233891 to AA234338, and AA41685 to AA41774 represent polynucleotide and propagatide sequence given in the exemplification of the present
                                                                                                                                                                                                                                                                                                                                                                                                                 giggagatattaactititititititititititiccitiggiggaagcigcictagggagggg 648
                                                                                                                                                                                                                                                                                                                                                                                                                                                            201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    261
                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                             gtcccatccattaagccatcacttctggaagattaaagttgtcggacatggtgacagctg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gtcccatccattaagccatcacttctggaagattaaagttgtcggacatggtgacagctg
                                                                                                                                                                                                                                                                                                      Length 4650
                                                                                                                                                                                                                                                                                                                                          .;
;
                                                                                                                                                                                                                                           0 other;
                                                                                                                                                                                                                                                                                                                                        18; Indels
                                                                                                                                                                                                                                                                                                  DB 20;
                                                                                                                                                                                                                                                                                                  Score 2033.8;
Pred. No. 0;
1; Mismatches
                                                                                                                                                                                                                                                                                                                                        1;
                                                                                                                                                                                                                                                                                                    89.8%;
                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                      Local Similarity
nes 2066; Conserv
                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                        Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                     589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                649
                                                                                                                                                                                                                                                                                                                                                                                 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                            82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            69/
```

셤 ð a òζ 요

ð

셤

ò

ctctcagaggtagcaaaggaacatattgggaaggagggatccggggctgtaggctttgtgc

1342

ò

```
1401
                                                                                                                                                                                                                                                                                                                                         1488
                                                                                                                                                                                                                                                                                                                                                                                         1041
                                                                                                                                                                                                                                                                                                                                                                                                                        1101
                                                                                                                                                                                                                                                                                                                                                                                                                                                        1161
                                                                                                                                                                                                                                                                                                                                                                        1548
                                                                                                                                                                                                                                                                                                         1428
                                                                             1008
                                                                                                             1068
                                                                                                                                                                                                           1248
                                                                                                                                                                                                                                           1308
                                                                                                                                                                                                                           741
                                                                                                                                                                                                                                                                                          861
                                                                                                                             561
                                                                                                                                                             621
                                                                                                                                                                                           681
                                              948
                                                              441
                                                                                              501
321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      agatgtacactcagagagtacagcaaatcttagcttcccataaccccacaaagcctatat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1189 cacctactcttgacaagctcgctgccgaaggagttaaactggagaactactatgtccagc
                                                                                                                                                                                                                            ctatttgcacaccatccaggagtcagtttattactggaaagtatcagatacacacgggac
                                                                                                                                                                                                                                            ctatttgcacaccatccaggagtcagtttattactggaaagtatcagatacacaccggac
                                                                                                                                                                                                                                                                                                                                   ttttatatattgcctatcaagctgttcattcaccactgcaagctcctggcaggtatttcg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ttttatatactgcctatcaagctgttcattcaccactgcaagctcctggcaggtatttcg
cttatggttacctgtcctggggccaggccttagaagaggaggaagaaggggccttactag
                                                                                                                            cacctactcttgacaagctcgctgccgaaggagttaaactggagaactactatgtccagc
                                                                                                                                                                                                                                                           ttcaacattctatcataagacctacccaacccaactgtttacctctggacaatgccaccc
                                                                                                                                                                                                                                                                                            tacctcagaaactgaaggaggttggatattcaacgcatatggtcggaaaatggcacttgg
                                                                                                                                                                                                                                                                                                                           gtttttacagaaaagaatgcatgcccaccagaagaggatttgatacctttttggttccc
                                         gaatggctcccaggggctgtgcgggggatccgcctccgccttctccacaggcctgtgtct
                                                                gtcctggaaagatgctagcaatggggggggctggcaggattctggatcctctgcctca
                                                                              gtcctggaaagatgctagcaatgggggggcgctggcaggattctggatcctctgcctcca
                                                                                                              1009 cttatggttacctgtcctggggccaggccttagaagaggaggaagaagggggccttactag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1849
                                                                                                                                                                                                                                             1249
                                                                                                                                                                                                                                                                           1309
                                                                                                                                                                                                                                                                                                            1369
                                                                                                                                                                                                                                                                                                                                           1429
                                                                                                                                                                                                                                                                                                                                                                                                          1549
                                                                                                                                                                                                                                                                                                                                                                                                                          1042
                                                                                                                                                                                                                                                                                                                                                                                                                                         1609
                                                                                                                                                                                                                                                                                                                                                                                                                                                         1102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1669
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1282
                                                                                                                               502
                                                                                                                                              1069
                                                                                                                                                               562
                                                                                                                                                                             1129
                                                                                                                                                                                                                             682
                                                                                                                                                                                                                                                                                                                            862
                                                                                                                                                                                                                                                                                                                                                           922
                                                                                                                                                                                                                                                                                                                                                                           1489
                                                                                                                                                                                                                                                                                                                                                                                          982
                 829
                                                889
                                                                                                                                                                                                                                                            742
                                                                                                                                                                                                                                                                                            802
                                                                                                442
                                                                                                                                                                                              622
                                 322
                                                                382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                        á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        合
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                              g
                                                                                                                                                                             a
                                                                                                                                                                                              ò
                                                                                                                                                                                                           g
                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                           ò
                                               g
                                                                                셤
                                                                                                                                              a
                                                                                                                                                             οy
              셤
                                 ò
                                                                δ
                                                                                                ò
                                                                                                                               ò
```

ř						
		•				
					5	
					•	i
						-

```
ద
                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                     g
  õ
                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                         å
                                                                                                                                                                                                                           2447
                                                                                                                                                                                                                                                                                            2060
                                                                                                                 2207
                                                                                                                              1701
                                                                                                                                                        1761
                                                                                                                                                                     2327
                                                                                                                                                                                                                                       2001
                                                                                                                                                                                                             1881
                                                                                                                                                                                                                                                                                                       2568 agaaaaagaagccaagcaaaaatcaggctgagaaaaagcaaaagaaaagcaaaaaaaga 2627
                                                                                                                                                                                                                                                                                                                                                                                                                         nootropic; neuroprotective; antibacterial;
                                             1909 ctctcagaggtagcaaaggaacatattgggaaggaggagggatcgtaggctttgtgc
                   gctatgatatctgggagaccataagtgaggtcttcgctcaccccgagtagatattttgc
                                                                                      gctatgatatctgggagaccataagtgagggtcttcgctcaccccgagtagatattttgc
                                                                                                                                    ggcacaatgaacggatcacctcgtcaactggcaaaagtgtatggcttttcaacatcacag
                                                                                                                                                                                                             ccgacccatatgagagggtggacctatctaacaggtatccaggaatcgtgaagaagctcc
                                                                                                                                                                                                                   agaaaaagaagccaagcaaaaatcaggctgagaaaaagcaaaagaaaagc-aaaaaaga
                                                                                                                             tctggaacactgcaatccagtcagccatcagagtgcagcactggaaattgcttacaggaa
                                                                                                                                                        atcctggctacagcgactgggtccccctcagtcttcagcaacctgggaccgaaccggt
                                                                                                                                                                                                                                                                                                                                                                                                                               antiinflammatory; vasotropic; antitumor; gene therapy; human; ss
                                                                                                                                                                                                                                                                                                                           2628 agaagaaacagcagaaagcagtctcaggtaaaccagcaaatttggct 2674
                                                                                                                                                                                                                                                                                                                      agaagaaacagcagaaagcagtctcaggttcaacttgccattcaggt
                                                                                                                                                                                                                                                                                                                                                                                                           sulfatase polypeptide coding sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
1..1800
/*tag= a
/product= "23553 sulfatase"
                                                                                                                                                                                                                                                                                                                                                                   CDNA; 1800
                                                                                                                                                                                                                                                                                                                                                                                                                      Sulfatase; 23553 sulfatase;
                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                   AAH46867 standard;
                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                           Human 23553
                                                                                                                                                                                                                                                                                                                                                                                              25-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                 AAH46867;
                                                                                                                                                                                                                                                                                                                                                      Ŋ
                                                                                                                                                                                                                                                   2448
                                                                                                                                                                                                                                                                                                                     2061
                                                                                     2089
                                                                                                                             1642
                                                                                                                                          2208
                                                                                                                                                                                                            1822
                                                                        1522
                                                                                                                                                       1702
                                                                                                                                                                                  1762
                                                                                                                                                                                                                         2388
                                                                                                                                                                                                                                       1882
                                                                                                                                                                                                                                                                 1942
                                                                                                                                                                                                                                                                              2508
                                                                                                                                                                                                                                                                                                                                                            AAH46867
                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                     요
                                                                                                               셤
                                                                                                                             ð
                                                                                                                                          윱
                                                                                                                                                                   셤
                                                                                                                                                                                             a
                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                 a
                                셤
                                                          g
                                                                         à
                                                                                                   ð
                                                                                                                                                       õ
                                                                                                                                                                                 ò
                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                 ò
                    ð
```

```
and Pick disease, spinocerebellar degenerations including spinocerebellar ataxias including Friedreich ataxia, and ataxia telangiectasia, degenerative diseases affecting motor neurons including amyotrophic lateral sclerosis, inborn errors of metabolism such as leukodystrophies, toxic and acquired metabolic diseases, including vitamin deficiencies, and neurocutaneous syndromes (phakamatoses) including neurofibromotosis. The present sequence represents a human 23553 sulfatase polypeptide
                                                                                                                                                                                                                                                                                                                                                                                                                       polypeptides and polynucleotides. The sulfatase genes and polypeptides are useful for treating disorders involving the brain such as cerebrovascular diseases, infections such as acute meningitis, demyelinating diseases including multiple sclerosis, degenerative diseases affecting the cerebral cortex including Alzheimer's disease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    623
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          743
                                                                                                                                                                                                                                                                            diagnosing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        324 atggeteceaggggetgtgeggggeatecgeetecgeettetecaeaggeetgtgtetgt 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09
                                                                                                                                                                                                                                                                     Novel human sulfatase polypeptides useful for treating and diagnosin sulfatase-related disorders such as cerebrovascular diseases, acute meningitis, multiple sclerosis, degenerative diseases and tumor
                                                                                                                                                                                                                                                                                                                                                                                                       invention provides 22438, 23553, 25278 or 26212 human sulfatase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 atggeteceaggggetgtgeggggeateegeeteegeettetecacaggeetgtgtetgt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              384 cctggaaagatgctagcaatggggggggcgctggcaggattctggatcctctgcctcact
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        cctggaaagatgctagcaatggggggggctggcaggattctggatcctcgctcctcact
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  tatggttacctgtcctggggccaggccttagaagaggaagaagaaggggccttactagct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          atttgcacaccatccaggagtcagtttattactggaaagtatcagatacacaccggactt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 1800;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4; Indels
                                                                                                                                                                  Ē
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 1800 BP; 534 A; 434 C; 423 G; 409 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 22;
                                                                                                                                                                       Ä
                                                                                                                                                                     Rudolph-Owen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 1781.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No.
                                                                                                                                                                                                                                                                                                                                                           2; Fig 15; 180pp; English
                                                                                                                                                                  Glucksmann MA, Williamson M,
                                                                                                                         (MILL-) MILLENNIUM PHARM INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       78.68;
99.78;
                                         31-JAN-2001; 2001WO-US03266
                                                                                 2000US-0495823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 99.7
Matches 1795; Conservative
                                                                                                                                                                                                          WPI; 2001-476214/51.
P-PSDB; AAB85483.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    coding sequence
                                                                                 31-JAN-2000;
02-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                           Claim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               444
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           744
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          684
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

Ξ;

480

421

셤

WO200155411-A2

Ī						
			٠			
					,	
	~				•	
					-	4